985

National Institute of General Medical Sciences

ANNUAL REPORT FY 1985

NEWARK PUBLIC LIBRARY

JAN 30 1986

U. S. GOVT. PUBS.
DEPOSITORY 375

U.S. DEPARTMENT OF HEALTH
AND HUMAN SERVICES
Public Health Service
National Institutes of Health

HE 20.3451:

READEX FICHE best copy available

CONTENTS

THE DIRECTOR'S REPORT..... 1

Fiscal Tables..... 3

REPORT ON RESEARCH TRAINING..... 5

CELLULAR AND MOLECULAR BASIS OF DISEASE PROGRAM..... 7

Objectives..... 7
Organization and Staffing..... 8
Research Overview..... 8
1. Cellular Basis of Disease Section..... 8
2. Molecular Basis of Disease Section..... 13
Research Training..... 14
Research Highlights..... 17
Program Organization and Staff..... 49

GENETICS PROGRAM..... 50

Objectives..... 50
Organization and Staffing..... 50
Research Overview..... 51
Research Training..... 54
Contracts..... 55
Small Business Innovation Research..... 57
Research Highlights..... 58
A. New Methodologies in Genetics..... 58
B. Developmental Genetics..... 68
C. Replication, Recombination, and Repair..... 73
D. Extranuclear Inheritance..... 94
E. Control of Gene Expression and RNA Splicing..... 101
Program Organization and Staff..... 115

PHARMACOLOGICAL SCIENCES PROGRAM..... 116

Objectives..... 116
Organization and Staffing..... 116
Research Overview..... 116
Research Training..... 118
Research Highlights..... 119
Program Organization and Staff..... 138

BIOPHYSICS AND PHYSIOLOGICAL SCIENCES PROGRAM..... 139

Scope and Objectives..... 139
Research Overview..... 139
1. Biophysics Section..... 139
2. Physiological Sciences Section..... 140
Shared Instrumentation..... 141
Small Business Innovation Research..... 141
New Investigator Research Awards..... 142
Research Training..... 142
Research Highlights..... 145
Biophysics..... 145
Physiological Sciences..... 157
Program Organization and Staff..... 162

MINORITY ACCESS TO RESEARCH CAREERS PROGRAM..... 163

Objectives..... 163
Program Components..... 163
Organization and Staffing..... 164
Highlights..... 165
Evaluation..... 165
Supplemental Activities..... 166
Program Organization and Staff..... 168

THE DIRECTOR'S REPORT

Fiscal Year 1985 marked the implementation of a reorganization of the programs of the National Institute of General Medical Sciences (NIGMS). In order to fully realize the potential of scientific opportunities in the biophysical sciences, a new program area was established, the Biophysics and Physiological Sciences (BPS) Program. It replaced and incorporated the research and research training activities of the former Physiology and Biomedical Engineering Program, as well as those aspects of the other programs which emphasize biophysical studies. Dr. Marvin Cassman was appointed as the Director of this Program. This reorganization, while capitalizing on the rapid growth in the biophysical sciences, also allowed for a better balance among the various programs. In addition, a very large number of the Institute staff have responsibilities in more than one program area, thus enhancing the collaborative and cooperative efforts among the programs. This is particularly important in an Institute which supports basic research in the sciences which will ultimately be of importance to the missions of the other institutes of NIH. For it is clear that the research supported by NIGMS is a continuum which, for administrative neatness, must be divided into discrete program areas but which dovetail and flow into each other. Thus the collaboration among the staff is of great importance to the total enterprise.

The scientific opportunities in the biomedical area continue to increase, and the NIGMS remains dedicated to supporting primarily investigator-initiated research as well as multidisciplinary research training of the highest quality. This is best illustrated by citing the support that the Institute provided in the mid-1970's for the study of steroid receptors on the cell membrane by Michael Brown and Joseph Goldstein. This fundamental research led to specific studies of familial hypercholesterolemia and of coronary artery disease by these two scientists which have, in more recent years, been supported by the National Heart, Lung, and Blood Institute and culminated in the award of the 1985 Nobel Prize in Physiology or Medicine to Drs. Brown and Goldstein.

The uncertainties regarding the availability of funds for research grants which were in effect through all but the last months of FY 1985 provided a challenge to the staff in regard to the setting of funding priorities and in maintaining a calm and appropriate dialog with members of the scientific community. The excellence of the members of the staff and the helpful and sympathetic attitudes they displayed while still carrying out their oversight responsibilities showed professionalism of the highest order, and they deserve great praise.

The staff of the Office of Review Activities had a heavy burden during this fiscal year, as it did in FY 1984, since the vast majority of the active training grants were up for renewal in FY 1985. In spite of heavy workloads, the staff was able to carry out careful reviews regarding the quality of research training at those institutions applying for initial or renewal grants. Since the number of awards that can be made is quite limited, these reviews were of extreme importance in the final decisions.

The Institute continues to support certain resources which provide the scientific community with needed materials or data to enhance research endeavors. Two research contracts, the Nucleic Acid Sequence Data Bank (GenBank®) and the Human Mutant Cell Repository, serve this purpose and are carrying out their functions in this regard very well. Particular thanks are extended to Dr. Christine Carrico,

who has served so well as project officer of GenBank® for the past three years and who will now relinquish this role to Dr. James Cassatt.

PERSONNEL CHANGES

Appointments

Dr. Christine Carrico, Director, Pharmacological Sciences (PS) Program
Dr. Marvin Cassman, Director, Biophysics and Physiological Sciences (BPS) Program
Dr. Judith Greenberg, Deputy Director, Genetics Program
Dr. Bert Shapiro, Deputy Director, Cellular and Molecular Basis of Disease (CMBD) Program
Dr. James Cassatt, Chief, Biophysics Section, BPS Program
Dr. Lee Van Lenten, Chief, Physiological Sciences Section, BPS Program
Dr. Warren Jones, Chief, Molecular Basis of Disease Section, CMBD Program

Additions to Staff

Dr. Marion Zatz, CMBD Program
Dr. Loré Anne McNicol, CMBD Program
Dr. Jane Peterson, Genetics Program
Dr. Janet Newburgh, PS Program
Dr. Rodney Ulane, Office of Review Activities (ORA)
Dr. Bruce Wetzel, ORA
Ms. Linda Engel, ORA

Retirements

Dr. Emilie A. Black, Assistant Director for Clinical Research
Dr. George Woolley, Genetics Program
Dr. Dorothea Starbuck Miller, Genetics Program
Dr. Vincent Price, CMBD Program
Dr. James Gilliam, PS Program
Dr. William Taylor, BPS Program

Departures

Dr. Sara Gardner, Director, PS Program
Dr. David Beck, Genetics Program
Dr. Anthony Demsey, ORA
Dr. Harriet Gordon, ORA

National Advisory General Medical Sciences Council

Newly appointed members are:

Dr. Norton B. Gilula
Dr. Harlyn O. Halvorson
Dr. Gerald Kanter
Dr. Mary Lou Pardue
Dr. Theodore Sherrod
Ms. Barbara Gill

NATIONAL INSTITUTE OF GENERAL MEDICAL SCIENCES
FY 1985 ACTUAL
(Dollars in Thousands)

	CHBD		GEN		PS		PBRE		MARC		RMPS		TOTAL	
	No.	Amount	No.	Amount	No.	Amount	No.	Amount	No.	Amount	No.	Amount	No.	Amount
RESEARCH GRANTS														
Research Projects														
New	161	$17,768	160	$19,840	76	$8,492	75	$6,863					472	$52,963
Competing	181	26,851	198	33,642	79	12,654	45	7,032					503	80,179
Supplemental	6	518	2	133	1	9	1	47					10	707
Subtotal Competing	358	45,137	360	53,615	136	21,155	121	13,930					987	133,845
Admin. Supp.	37	1,772	26	983	0	187	4	56					75	2,998
Noncompeting	725	93,221	783	114,899	283	34,427	116	17,443					1,907	259,990
Subtotal Regular	1,075	140,130	1,143	169,497	439	55,769	237	31,437					2,894	396,833
Research Centers														
New													0	0
Competing													0	0
Supplemental	1	687											1	687
Subtotal Competing	1	689											1	689
Admin. Supp.	1	74					1	68					2	142
Noncompeting			4	3,159	4	2,128	6	3,117					14	8,404
Subtotal Centers	1	763	4	3,159	4	2,128	6	3,185					15	9,235
Other Research														
Res. Career Awards														
Noncompeting	2	66	6	200	1	32	2	76					11	374
Subtotal RCA	2	66	6	200	1	32	2	76					11	374
Res. Car. Dev. Awds.														
New							1	51					1	51
Noncompeting					2	91	2	93					4	184
Subtotal RCDA					2	91	3	144					5	235
Total RCP														
New							1	51					1	51
Noncompeting	2	66	6	200	3	123	4	169					15	558
Subtotal RCP	2	66	6	200	3	123	5	220					16	609
Scientific Evaluation								2,142					0	2,142
(DRG)								(1,070)					0	(1,070)
(BR)							4	(1,072)					4	(1,072)
Other (R13)														
New	12	24	6	9	2	3	2	4					22	40
Competing	1	110											1	110
Supplemental													0	0
Subtotal Competing	13	134	6	9	2	3	2	4					23	150
Noncompeting			1	42									1	42
Subtotal	13	134	7	51	2	3	2	4					24	192
Total Other														
New	12	24	6	9	2	3	5	490					25	526
Competing	1	110											1	110
Supplemental													0	0
Subtotal Competing	13	134	6	9	2	3	5	490					26	636
Noncompeting	2	66	7	242	3	123	6	1,876					18	2,307
Subtotal Other	15	200	13	251	5	126	11	2,366					44	2,943
TOTAL RESEARCH GRANTS	1,091	141,093	1,160	172,907	448	58,023	254	36,988					2,953	409,011
NRSA Individual														
New	58	1,182	111	2,260	34	650	11	260	19	430			233	4,782
Competing			1	26					2	38			3	64
Supplemental													0	0
Subtotal Competing	58	1,182	112	2,286	34	650	11	260	21	468			236	4,846
Noncompeting	65	1,458	143	3,301	18	380	2	51	38	632			266	5,822
Subtotal Indiv.	123	2,640	255	5,587	52	1,030	13	311	59	1,100			502	10,668
NRSA Institutional														
New	2	139					6	257	5	218			13	614
Competing	40	10,917	1	42	17	1,753	22	8,681	14	1,556			94	22,949
Supplemental	1	51											1	51
Subtotal Competing	43	11,107	1	42	17	1,753	28	8,938	19	1,774			108	23,614
Noncompeting	42	7,274	12	1,285	28	2,916	30	7,097	41	4,867			153	23,439
Subtotal Instit.	85	18,381	13	1,327	45	4,669	58	16,035	60	6,641			261	47,053
TOTAL NRSA	208	21,021	268	6,914	97	5,699	71	16,346	119	7,741			763	57,721
TOTAL GRANTS	1,299	162,114	1,428	179,821	545	63,722	325	53,334	119	7,741			3,716	466,732
R & D Contracts														
New	1	40	2	455									3	495
Renewal	1	10											1	10
Continuation			1	729									1	729
Subtotal, Contracts	2	50	3	1,184									5	1,234
Intramural Res.											22	616		616
Res. Mgmt. & Support											139	12,393		12,393
(Mgmt. Fund)												(4,313)		
(General Exp.)												(216)		
(Program Eval.)												(1,295)		
(Consultant)												(89)		
Total RMS												13,009		13,009
TOTAL NIGMS	1,301	162,164	1,431	181,005	545	63,722	325	53,334	119	7,741		13,009	3,721	480,975

NATIONAL INSTITUTE OF GENERAL MEDICAL SCIENCES
FY 1985
TRAINING PROGRAMS

	CMBD			GENETICS			PHARM. SCI.			PBME			MARC			TOTAL		
	FTE	NO.	AMOUNT	FTE	NO.	AMOUNT	FTE	NO.	AMOUNT	FTE	NO.	AMOUNT	FTE	NO.	AMOUNT	FTE	NO.	AMOUNT
INDIVIDUAL - PREDOCTORAL:																		
PREDOCTORAL													37	37	$504	37	37	$504
MARC FACULTY FELLOW													12	12	334	12	12	334
MARC VISITING SCIENTIST													1	1	29	1	1	29
TOTAL													50	50	867	50	50	867
INDIVIDUAL - POSTDOCTORAL:																		
POSTDOCTORAL	122	122	$2,614	255	255	$5,587	52	52	$1,030	11	11	$245	0	0	0	440	440	9,476
SENIOR FELLOWS	1	1	26							2	2	66				3	3	92
MARC FACULTY FELLOW													9	9	233	9	9	233
TOTAL	123	123	2,640	255	255	5,587	52	52	1,030	13	13	311	9	9	233	452	452	9,801
TOTAL, INDIVIDUAL	123	123	2,640	255	255	5,587	52	52	1,030	13	13	311	59	59	1,100	502	502	10,668
INSTITUTIONAL - PREDOCTORAL:																		
PREDOCTORAL	1,204	81	18,017	0	0	0	273	33	3,679	185	23	2,748	28	2	316	1,690	139	24,760
MEDICAL SCIENTIST										683	25	12,597				683	25	12,597
MARC UNDERGRADUATE													387	53	6,075	387	53	6,075
SHORT-TERM TRAINING										10	2	87				10	2	87
MARC VISITING PROFESSORS													0	4	145	0	4	145
TOTAL	1,204	81	18,017	0	0	0	273	33	3,679	878	50	15,432	415	59	6,536	2,770	223	43,664
INSTITUTIONAL - POSTDOCTORAL:																		
POSTDOCTORAL	13	4	364	51	13	1,327	36	12	990	20	8	603	4	1	105	124	38	3,389
TOTAL INSTITUTIONAL	1,217	85	18,381	51	13	1,327	309	45	4,669	898	58	16,035	419	60	6,641	2,894	261	47,053
TOTAL	1,340	208	21,021	306	268	6,914	361	97	5,699	911	71	16,346	478	119	7,741	3,396	763	57,721

Cooperative Agreements with other Institutes

MARC PREDOCTORAL	1	1	$10
MARC FACULTY FELLOWS	8	8	131
TOTAL	9	9	141

RESEARCH TRAINING

Since its establishment, the NIGMS has had, as one of its principal responsibilities, the support of research training in the biomedical sciences. In fact, through a variety of programs, this Institute provides support for two-thirds of all the predoctoral trainees supported by the NIH. In the future, these trainees are expected to be among the leading scientists who will fill the nation's health research needs and thus carry out the mission of the NIGMS and the NIH.

All training support at the NIH is provided under extension and renewal of the authority of the National Research Service Award Act of 1974. Under this legislation, the NIGMS supports predoctoral graduate training through multidisciplinary and multi-departmental research training grants. The intent of these training grant awards is to allow broad and fundamental research training in the basic biomedical sciences. Within each of the four multidisciplinary clusters (cellular and molecular biology, genetics, pharmacological sciences, and systems and integrative biology) are a wide variety of programs at universities and medical schools throughout the country. A fifth program, the Medical Scientist Training Program (MSTP), provides support for students seeking the combined M.D.-Ph.D. degree at outstanding medical schools in the United States. This program is designed to train an exceptional breed of scientist who will play a role in easing the national shortage of clinical investigators.

Postdoctoral support is emphasized somewhat less by the NIGMS, and is provided primarily through individual fellowships. This type of award serves several basic science areas and is appropriate when the postdoctoral scientist is working intensely with an established investigator on a particular research problem, with little need for an organized training program. The NIGMS does, however, award a small number of postdoctoral research training grants in the more clinically related areas of research training, and emphasizes the selection of M.D. degree holders as trainees.

The Minority Access to Research Careers (MARC) Program is a unique research training program administered by the NIGMS in collaboration with other institutes of the NIH. Support is provided for research training in health-related areas for faculty and undergraduate students at minority institutions throughout the country. The most outstanding of these students are subsequently supported by individual fellowship awards to complete their Ph.D. degree at universities of their choice. A more detailed description of this, as well as the other training programs, is found in the MARC program section of this report.

The NIGMS research training budget in FY 1984 was equal to the FY 1983 level of $48.8 million. In FY 1985, the research training budget was increased to $57.7 million to permit a substantial increase in the level of stipends provided. The new predoctoral stipend is $6,552; the new postdoctoral stipend starts at $15,996. In each of the last two years, the number of trainees receiving support decreased. A major feature in the allocation of the research training budgets during recent years has been the sharp increase in the tuition category, resulting in a further decrease in the trainee numbers. Although actual tuition varies widely from school to school, in the last

eight years, average tuition costs have increased from $2,700 to $6,200.

In spite of these trainee reductions, careful attention has been given to new applications from institutions and training programs having no previous NIGMS support. During the past two years, the Institute was able to make 12 awards to new programs: four in the MARC undergraduate program, seven in the predoctoral program, and one in the postdoctoral program.

For a variety of reasons, including legislative changes, recommendations by the National Academy of Sciences, budgetary constraints, and, in particular, the steadily increasing costs per trainee--the total number of students supported by the NIGMS in the regular predoctoral program has decreased from 4,500 per year 12 years ago to less than 1,700 per year in FY 1985. The other predoctoral program, the MSTP, has been more constant, supporting about 680 students per year for each of the past five years. The long-range impact of this overall downward trend cannot be fully predicted, but the Institute is aware of concerns expressed by members of the scientific community that further decreases could have a deleterious effect on the viability of the research training programs at some institutions since a critical mass at each institution is necessary for success. Further Institute planning efforts will factor in these concerns.

Since the enactment of the National Research Service Award Act in 1974, the overall NIGMS research training program has utilized three funding mechanisms: predoctoral institutional multidisciplinary training grants, individual postdoctoral fellowships, and postdoctoral institutional training grants. The National Advisory General Medical Sciences Council had previously recommended that the Institute support research training in that priority order and reaffirmed that view this year.

CELLULAR AND MOLECULAR BASIS OF DISEASE PROGRAM

OBJECTIVES

The cell is the fundamental biological unit of all living organisms. It holds the capacity for generating, from simple exogenous sources, all its macromolecules, energy, and complex organization and structure. The Cellular and Molecular Basis of Disease Program (CMBD) is concerned with basic questions about the function of cells and their components. This research and training support encompasses a wide range of scientific efforts, from cell-cell communication, cell adhesion, and development, to cellular metabolism, energetics, enzyme mechanisms, and the molecular structure of cells. Ranging as it does, from the molecular to the multicellular, the program is broad in scope, and supports work utilizing a wide variety of technologies, instruments, and models.

A major area showing recent expansion is that of cell-cell signalling and cellular interactions. Many interesting signal compounds have been studied intensively. These include the lymphokines, growth factors, hormones, and mitogens. Although other categorical institutes may provide significant support of research into the agents, e.g. NIAID for lymphokines, NIADDKD for hormones, NCI for mitogens and growth factors, the agents and their targets are rarely confined to one system. For example, the lymphokines not only promote lymphocyte growth, but also are important in thermoregulation. Another example is the peptides now being intensively studied as signal molecules in the central nervous system that were originally identified as hormones controlling digestive physiology. Lastly, CMBD, and NIGMS in general, supports research using a very wide range of models. Cells or processes of great interest to categorical institutes, such as transport in erythrocytes, may be excellent models for attacking fundamental questions. Thus this program supports many projects into the mechanism of selectivity and transport kinetics of erythrocyte chloride channels (see under 'Highlights'). Although many of the projects use prokaryotes as models, or even completely synthetic systems, such as the study of synthetic antibiotic channels in black lipid membranes, many others use vertebrate, or even human material. This trend is due to the fact that there have been major strides in the use of human cells. Tissue culture methods are always improving, and mammalian cellular genetics and molecular biology are now better understood. Human genes can be cloned and expressed; loci can be mapped much more easily; mammalian viruses are beginning to prove as useful for the study of complex cells as bacterial viruses have been for bacteria. At the same time, on the other end of the scale, it is becoming easier to alter proteins in specific ways. This has opened up a major biochemical area: the structure-function relationship of proteins.

Expansion of the use of higher organisms, as well as of altered peptides, only begins to tell the story of modern cell and molecular biology. In an era of increased specialization, the opposite trend is observed in the areas supported by the CMBD program. Investigators are called on more and more to utilize a spectrum of techniques and approaches to attack biomedical problems. The study of microtubules may require isolation of the tubulin gene, immunoelectron microscopy (itself a hybrid field), tissue culture, photoactivatable crosslinking, protein purification, and cell growth control. The diffusion of new techniques throughout the biomedical research community is increasingly rapid. The oppor-

tunities for sophisticated investigation have never been greater, and the CMBD program supports the highest quality cellular and molecular research and training.

The program's specific objective is to gain the fullest knowledge about human cells to assist in the prevention, treatment, and cure of disease in man. Since living cells generally share fundamental biological, chemical, and physical processes and properties, the CMBD Program supports research using cell and molecular systems from a wide variety of species, tissues, and preparations. The basis of selection is the scientific merit of research using preparations best suited to solving general principles and problems. In pursuit of this objective, the program supports research ranging from precise physical and mathematical approaches in the study of enzyme catalyzed reactions to investigation of the cellular interactions during embryonic development. Results from this broad range of research are reported in many scientific journals, adding important general principles, techniques, and information which contribute directly and indirectly to our expanded knowledge about normal and abnormal cellular functions. It is this fundamental knowledge that undergirds the study of human function and disease.

ORGANIZATION AND STAFFING

The overall CMBD research program is administered by a Director and Deputy Director, and five other professional staff members. It is organized into two sections and subprogram areas. The professional staff is:

Charles A. Miller, Ph.D.	Director
Bert I. Shapiro, Ph.D.	Deputy Director

Cellular Basis of Disease Section:

Bert I. Shapiro, Ph.D.	Chief
Artrice V. Bader, Ph.D.	Program Administrator
Marion M. Zatz, Ph.D.	Program Administrator

Molecular Basis of Disease Section:

Warren C. Jones, Ph.D.	Chief
Lee Van Lenten, Ph.D.	Program Administrator

Research training is administered separately from the section organization. Drs. Miller, Bader and von Euler (Deputy Director, NIGMS) are responsible for training.

RESEARCH OVERVIEW

1. CELLULAR BASIS OF DISEASE SECTION

The emphasis of the Cellular Basis of Disease Section is on research on the cell and its subcellular components including basic research on all types of cells (i.e., prokaryotic and eukaryotic microorganisms), cells in tissue or

organ culture, isolated cells, such as blood cells, sperm, and ova, as well as research focused on specific cells and their function in excised and intact tissues and organisms. Questions and problems explored and pertaining to specific cell types or components are of a fundamental or general nature applicable to other cell types. This section includes all classes of basic cellular or subcellular research not expressly directed to the disease-oriented mission of a single categorical institute, as well as cell biology related to several disease or general pathological states. The full range of physical, chemical, and biological methods are employed throughout, and the development of new methods and techniques are provided.

This section includes programs in Cell Regulation, Differentiation and Growth; Molecular Immunobiology; Cell Organization, Motility and Division; and Membrane Structure and Function.

Cell Regulation, Differentiation and Growth

This subprogram supports a wide spectrum of investigations on normal cell functions that encompass many active areas of cell biology. About two thirds of the research areas center on two major themes--cell differentiation and the cell cycle. The other third supports a wide selection of research on diverse topics including cell adhesion and interaction, circadian rhythms, and the interaction of cell water with proteins and small molecules. Specifically, the major research areas covered are: growth, differentiation, and activation of all cell types; transmembrane signalling and receptor coupling with cellular response; genetic and microenvironmental regulation of differentiation; cell-cell interactions and the extracellular matrix; organelle developments, differentiation, and cycling, including mitochondria, lysosomes, golgi apparatus, and the endoplasmic reticulum; the biochemistry and regulation of the cell cycle; and cytoplasmic messengers of all types, including hormones, ions, cyclic nucleotides, prostaglandins, and leukotrienes.

This program is a diverse one, in which investigators use varied experimental systems as different as the diatom and cultured mammalian cells. While the active areas of research in this program span the topics mentioned above, there is an increasing emphasis on selected aspects. These newer and expanding research areas include: regulation of growth and the cell cycle by growth factors and their subsequent stimulation by second messengers, phospholipids, and phosphoproteins; the role of calmodulin and calcium in cell processes; complex cell interactions in differentiation and development, using multicellular, prokaryotic and mammalian systems; and the role of regulation of gene expression during cell activation and response.

Recent specific areas supported by this subprogram which indicate the opportunities in the field are several. One exciting area is that of stress or heat shock proteins. Several labs are making excellent progress in purifying stress proteins. These proteins play an important role in modulating protein synthesis, and thus are of great interest in the study of control of gene expression. Another, although extracellular, protein involved in growth control is platelet-derived growth factor (PDGF). Grantees are studying the PDGF-inducible c-myc gene, a gene which may play an important role in cell response to mitogens. These two examples demonstrate the opportunities and central importance in the study of growth control and other areas in this CMBD subprogram.

Molecular Immunobiology

Immunobiology is a basic medical science which has implications for many clinical areas including host resistance to infectious agents, immunodeficiency, allergic reactions, autoimmune diseases, and diagnosis and treatment of neoplasias. Within NIGMS, research in basic immunobiology transcends the major programmatic areas of cellular and molecular biology, genetics, and biophysical and physiological sciences. Research on fundamental cellular and molecular processes, as studied in experimental models using cells of the lymphoid system and their products, is therefore supported within an interdisciplinary subprogram of Molecular Immunobiology.

The emphasis of this program is on the study of basic cellular processes at the molecular level, using cells of the lymphoid system and their products as experimental models. Areas of immunobiology which are of interest include: protein structure/function relationships of the cytokines, immunoglobulins, MHC products, and complement; membrane structure and function, exocytosis and endocytosis, coupling of receptor-ligand binding and cell response, transmembrane signalling and membrane/cytoskeletal relationships; regulation of cell growth and the cell cycle by cytokines, growth factors, and hormones; intracellular mediators of cell activation such as calcium, cyclic nucleotides, phosphoproteins, phospholipids and protein kinases; the biochemical basis of cell interactions, including the role of the MHC products, lymphokines and lectins, cell motility, cell adhesion and the extracellular matrix; cell differentiation and maturation; regulation of gene activation, gene organization and reorganization, developmentally regulated genes, extended gene families, and the genetic basis of the above cellular processes.

This newly formed subprogram area has arisen in response to a growing convergence of research in cell biology and immunobiology. Many immunobiological systems serve as excellent models for exploring basic questions of cell interactions, cell growth and differentiation, and regulation of gene expression. This program is filling a need for interdisciplinary support of the rapidly expanding areas of lymphocyte receptor-ligand coupling and transmembrane signalling; regulation of the cell cycle by cytokines and lymphoid growth factors; and the evolutionary and developmental significance and function of "supergene families" encoding for the imunoglobulins, major histocompatibility antigens, T-cell receptor, and lymphocyte differentiation antigens (e.g., Thy-1. Lyt. 2)

Cell Organization, Motility and Division

This subprogram includes research on the chemical and physical nature of cytoskeletal and other structural elements (e.g., microtubules, microtubule-associated proteins, actin, myosin) involving their relationship and role in cell organization, as well as cell streaming, ciliar and flagellar motions; associated taxis and tropism; and movement of internal macromolecular structures, and their interrelationships to membranes and organelles in secretion, endocytosis and other functions. It also supports research on cell division centering around the mitotic apparatus; chromosome and associated macromolecule duplication; and movement, as well as the regulation of these processes.

This subprogram area includes grant supported research which emphasizes cell motility and the cytoskeleton. The research on cell motility involves studies

of motion of and within non-muscle cells. These basic research studies include both unicellular and multi-cellular organisms, cells in tissue or organ culture, and isolated cells such as sperm and ova, both fertilized and unfertilized.

In addition to the movement of cells and cellular organelles, research pertaining to the microscopic, submicroscopic and molecular organization of cells and cell organelles is included in this subarea. Regarding cell movement, the earlier years of this subprogram reflected more interest along the lines of gross cellular movements including such activities as protoplasmic streaming and the bending of cilia and flagella. Although gross movement continues to be of great interest and still reflects the major theme of much of the work supported here, emphasis has over the past years shifted from gross anatomical and physiological studies to molecular biology studies employing many of the new technologies. Recombinant DNA technologies, immunological techniques, polyclonal and monoclonal antibodies, electron microscopy and high resolution fluorescence microscopy are now standard tools.

Also, fast emerging in this area is the use of computers and imaging processing technologies to answer unresolved questions and amplify existing answers regarding cellular and cellular organelle movements. The larger portion of the work supported here involves studies of the filaments comprising the cytoskeleton: microtubules, intermediate filaments and microfilaments; and the proteins associated with motile structures, including microtubular associated proteins, tubulin, myosin, paramyosin, actomyosin, and various protein kinases. Because of the role of actomyosin complexes in muscle cells this protein system serves as a model for elucidating movement in non-muscle cells. Although there is evidence for a possible role for actomyosin in non-muscle cells there is also evidence indicating a similar role for the dynein and dynein-ATPase systems found in non-muscle cells. The dynein-ATPase model is being used to explain organelle movements such as the mitotic spindle and flagellar arm movement. Many researchers are focusing their efforts to study the dynein moiety itself while others are investigating its role as a counter-part mechanism of the actomyosin system so thoroughly characterized in muscle cells.

Motility studies involve control of shape associated with normal as well as transformed cells. Studies of the role of ions, principally Ca++ ions and sulfhydryl groups, in cell movement and cell division are included here and calcium involvement in the regulation of motility continues to be a recurring theme in many of the projects. The biological role of intracellular Ca++, as well as the function of calcium and calmodulin and related proteins are of great interest. Their function is being defined in microtubular initiation and elongation studies, as well as regulation of microtubules and microfilaments; actomyosin interactions; ciliary motion and mitotic spindle assembly and function.

Cell division studies which center around the mitotic spindle, the movement of chromosomes and associated cellular proteins involved in the division process, are also included in this program area. Research involving the mitotic spindle includes studies relating to the composition of this structure, its isolation, the movement of related structures within this apparatus and its relationship to the cell membrane. Also a rapidly emerging area reflected in this subprogram is the interaction of cytoskeleton structures, and the cellular membrane during movement and division. Representative of these are the cytoskeleton-membrane interactions involving linker proteins such as vinculin and talin which are involved in attachment of the cytoskeleton to the membrane.

As new questions arise and methods for answering them evolve, one continously sees an overlap of scientific disciplines and, consequently, program and subprogram area overlap. In the case of this subprogram, there are many overlapping areas with the genetics program such as studies related to heat shock proteins, multi-gene families of tubulin, genetic analysis, studies of dynein arms in flagella, and translational and post translational activities during various biological events such as flagellar replication. Within the CMBD program, this group of grants overlaps with both the cell replication, differentiation and growth area and the membrane structure and function subprogram.

Membrane Structure and Function

This subprogram deals with all aspects of the biology and chemistry of membranes. The emphasis is on the plasma membrane of cells, but a significant fraction of the effort is devoted to intracellular membranes. The spectrum of research supported represents the wide range of approaches to the role of membranes. It has become increasingly clear that membranes play a central role in cell physiology, and have unique properties associated with the lipid phase and lipid-water interface. Among the topics included in this subprogram are the chemistry (including physical chemistry) of membrane proteins and lipids, including the establishment and maintenance of a bilayer, lipid turnover, and protein insertion. A large group of grants are concerned with various aspects of membrane transport, from passive diffusion to the energetics of active transport. Another major fraction of grants is concerned with membrane biogenesis, cycling, secretion and exocytosis, as well as endocytosis. The program includes many grants studying or utilizing artificial membranes of several types, including black lipid membranes and vesicles. Some of these projects also address the mechanism of selectivity and transport of ionophores. Membrane receptors are extensively studied, and NIGMS supports a significant fraction of these studies, especially when the questions concern basic, general receptor properties. Related to the grants on the physical and chemical stability of bilayers or membranes in general, are several grants studying membrane disruptions by viral or toxin lysis, or complement fixation.

The membrane subprogram is the largest in CMBD and also the most rapidly expanding. This reflects the growing realization of the importance of membranes, the opening up of exciting new lines of membrane research, and technological advances in the study of membrane proteins. The importance of membranes has long been recognized, but it has now become apparent that the membrane plays a crucial role in such processes as growth control, hormone action, protein processing, and cell energetics. Three of the exciting lines of research which have opened up are: the understanding of the mechanisms of protein insertion into membranes and protein translocation; the emerging picture of the pathways of membrane cycling; and the chemistry of membrane receptors. The number of grants devoted to the mechanism of the insertion of proteins, including the nature and role of leader sequences, is growing exponentially. The two areas of membrane cycling in which there have been major findings are in both membrane and membrane protein processing by the Golgi, and the internalization of membrane receptors (often with the signal molecules) from the coated pits. Membrane receptor chemistry has been given a significant boost by the same technological advances which have affected the entire membrane (and, indeed, CMBD) program, namely recombinant DNA technology. This technology has allowed purification and sequencing of rare membrane proteins of great interest, in-

cluding T-cell receptors, sodium channels, and glucose transporters. It is apparent that we are seeing just the beginnings of a field which promises to expand rapidly in the near future.

2. MOLECULAR BASIS OF DISEASE SECTION

The Molecular Basis of Disease Section supports basic research on enzymology and regulatory processes in intermediary metabolism; structure and function of the enzyme active site, including the involvement of coenzymes and metal ions; and characterization of the energy transducing apparatus. Essentially, most areas of basic biochemistry, and much of physical biochemistry of general biomedical significance are included. Support is provided for many studies aimed at basic principles, regardless of the particular experimental system.

Molecular Basis of Enzyme Catalysis and Regulation

The major emphasis of this program area is the analysis of intermediary metabolism on a molecular level. This ranges from a very narrow focus, such as probes of enzyme active sites, to broader constructs such as overall regulation of metabolic pathways. In general, the grants are concerned with events at the enzyme level. Included in this subprogram are studies of the structure of catalytic and regulatory sites of enzymes, employing physical, chemical or genetic methods. A large moiety consists of grants directed to understanding enzyme mechanisms, including kinetics studies, investigations of catalytic intermediates, and research using model compounds. A recently expanding area is that of enzymatic regulation, at both the protein and gene level.

One of the most exciting areas in this subprogram is that of mechanistic enzymology--investigations of the modes of catalysis and regulation of a broad range of enzymes. These investigations seek to establish in molecular terms the factors that govern the speed and specificity of enzymatic catalytic processes as well as the exquisite regulation of enzymes that contributes to the capacity of cells to respond to a changing environment. Using a combination of site-directed mutagenesis with very sensitive physical techniques, such as NMR and high-speed crystallography, biochemists can now explore, in detail, the conditions for optimal catalysis. Studies on one enzyme are now being extended to related enzymes, often of greater complexity. Mechanistic enzymology promises to make contributions to a wide variety of fields. It will provide a deeper understanding of the genetic defects which result in faulty cellular enzymes. It may well lead to rational drug design, and, in fact, promises to be of significant value to the assortment of commercially valuable endeavors known as biotechnology.

Glycoconjugates

A portion of the research projects supported by the Cellular and Molecular Basis of Disease Program focuses on the study of glycoconjugate structure, function, and metabolism, with emphasis directed towards the carbohydrate portion of the macromolecule. More targeted studies involving glycoconjugates can be found in other grant portfolios funded by both NIGMS and other institutes.

Much of the excitement and many of the recent advances in the area have come from the studies concerning mammalian glycoprotein biosynthesis and function.

The distribution of grants in this subset reflects this emphasis. Several of the supported projects focus on the study of dolichol metabolism, lipid-oligosaccharide biosynthesis and its regulation, the conversion of proteins to glycoproteins, or postsynthetic processing. Others address some of the components required in the metabolism of glycoproteins such as the glycosyltransferases or the N- and O-glycosidases, or the mechanism by which glycoproteins are targeted to particular subcellular organelles. While the majority of the grants in the portfolio concentrate on the study of mammalian glycoprotein metabolism and function, support is provided for projects whose primary emphasis is methodology development, structure determination, or the study of non-mammalian glycoconjugates, e.g., crustacean and fungal chitin, yeast mannoprotein, fungal glycoproteins, corona virus glycoproteins.

Bioenergetics

The Bioenergetics subprogram supports investigations of the processes by which energy is captured and made available to drive all cellular functions. Most of these investigations seek to delineate the structural and mechanistic details characteristic of mitochondrial oxidative phosphorylation and electron transport, as well as the related processes of photosynthesis and photophosphorylation in bacteria and other organisms.

No set of processes are more central to life than bioenergetics. The universality of many of the processes makes it advantageous to investigate mitochondrial energetics using lower eukaryotic mitochondria, or looking at some simpler models, such as bacteria. One area of major advancement is in our understanding of the F_1 ATPase of mitochondria. Other significant advances have been made in analyzing the details of electron transport in the cytochrome bc_1 complex. Despite the significance of bioenergetics, progress has been difficult because multicomponent, multienzyme complexes are involved, and the chemical events are extremely rapid. In addition, the nature of some of the molecular processes in electron transport is very different from most enzymatic catalysis. Recently, advances in biophysical techniques and in recombinant DNA technology has led to increasing sophistication in our understanding of the structure and function of the multitude of interdependent proteins that participate in the energy transduction process.

RESEARCH TRAINING

Support of research training at the predoctoral and postdoctoral levels remains an important activity within the program. About one-half of the Institute's major predoctoral research training activities and one-quarter of the NIGMS postdoctoral trainees and fellows are supported through the CMBD program.

Predoctoral research training grants are awarded in two related multi-disciplinary scientific areas entitled "Cellular and Molecular Biology" and "Genetics". These grants are awarded to universities and medical schools to assist in the support of integrated graduate degree programs, providing funds for stipends, tuition, and thesis research supplies for highly selected candidates pursuing studies for the Ph.D. degree. Students supported by the cell and molecular biology grants may seek degrees in biochemistry, biophysics, cell biology, chemistry, developmental biology, genetics, immunobiology, neurobiology, pathology,

physiology and other related disciplines. Students supported by the genetics grants may also receive degrees in a number of the disciplines mentioned but with highly focused attention to a central theme of genetics in the course of their didactic studies and thesis research projects. A few of these grant-supported programs emphasize indepth coursework and opportunities for thesis research in areas of population genetics or medical genetics.

In FY 1985, awards were provided to 47 cell and molecular biology programs for the support of 841 predoctoral trainees and 34 programs in genetics for support of 377 trainees. The overall funds provided for the 81 grants was $18,380,000. Forty-two of these grants were funded as competing continuation awards and only one was a new award--the latter for support for a highly regarded new program in cell and molecular biology. The 81 awards were to 53 universities throughout the country. Numerous faculty-scientists who are considered leaders in their fields of research, participate in these programs providing extensive interdepartmental, interdisciplinary research training opportunities. Increasing emphasis is generally placed on providing new approaches in the understanding of physical-chemical and quantitative analysis of a great variety of biological systems and problems. Significantly fewer programs and numbers of trainees were supported in FY 1985 than in FY 1984 due to sharply rising tuition costs and increases for trainee stipends within the funds available.

At the postdoctoral level, the CMBD program funded five research training grants in the area of "Pathobiology" for the support of 17 trainees. These programs select largely post-M.D. candidates who pursue indepth research problems under the guidance of outstanding scientists for a period of one to three years. Research opportunities may be selected from a host of cellular and molecular fields. Funds for this area were $470,000 in FY 1985.

Additionally, funds in the amount of $2,640,000 supported 123 individuals, largely post Ph.D.'s, via competitive individual fellowships to join the laboratory of an established scientist for a two or three year period. A large number of these fellows had earned the Ph.D. degree in chemistry and sought to gain advanced research training in an active area of biomedical science such as biochemistry, cell biology or molecular biology.

Many additional research training grants and fellowship applications were reviewed and recommended as highly worthy of support in the above areas but could not be considered for funding.

Research and Training Areas Supported by the Cellular and Molecular Basis of Disease Program

	No. of Grants*	Dollars (TC)
Cellular Basis of Disease Section		
Cell Regulation, Differentiation & Growth	85	10,273,328
Molecular Immunobiology	36	5,386,842
Cell Organization, Motility & Division	119	15,684,713
Membrane Structure & Function	312	43,167,566
Section Total	552	74, 512,449
Molecular Basis of Disease Section		
Enzyme Catalysis & Regulation	224	27,645,093
Glycoconjugates	34	4,076,829
Bioenergetics	74	10,378,334
Section Total	322	42,100,256
TOTAL RESEARCH SUPPORT	874	116,612,705

RESEARCH TRAINING SUPPORT

Category	No. of Trainees	No. of Grants*	Dollars (TD)
Institutional Predoctoral			
Cellular & Molecular Biology	841	47	12,706,499
Genetics	377	34	5,309,969
Total Predoctoral	1,218	81	18,016,368
Institutional Postdoctoral			
Basic Pathobiology	17	5	469,091
Individual Postdoctoral			
Cellular & Molecular Biology	123	123	2,640,000
Total Postdoctoral	140	128	3,109,091
TOTAL TRAINING SUPPORT	1,358		21,125,459

*active as of 09/01/85

RESEARCH HIGHLIGHTS

Presented below are some highlights of research supported by the CMBD program. Of course these represent a fraction of the excellent and exciting work done by grantees. These highlights were selected to underline the diversity of quality research funded by the program. Space limitations prevent inclusion of many highlights, and we try not to repeat general topics for several years.

What stands out in these highlights is the dramatic expansion in the use of recombinant DNA technology as a tool to attack many problems in cellular biology and biochemistry. For example, as will be discussed further, Dr. Frank Maley has been studying thymidilate synthase for several years, and the isolation of the enzyme from Lactobacillus, and its sequencing by classic protein chemistry techniques represents the culmination of years of effort. Subsequently, the enzymes from E. coli and phage were studied with the aid of recombinant DNA techniques, with more data obtained in a fraction of the time. This technology permitted isolation of the gene and expression in a vector under conditions which dictate overproduction. Much larger quantities of the enzyme are available via this strategy. Moreover, sequencing of the gene is far faster than direct protein sequencing. The nucleotide sequence obtained provides information not only on the final protein product, but also on the primary product of the translation. These peptides often include leader sequences which aid in membrane translocation and targeting. Other terminal moieties may be present only in the inactive proenzyme, to be activated subsequently under appropriate conditions. Therefore, knowledge of the primary translation product may provide us with additional information about the role and processing of the functional protein. Sequencing of the gene can also tell us about its regulation and evolution. By using the techniques of molecular biology to obtain information about protein biochemistry, scientists also learn more about molecular biology. The boundaries between the two fields shrink, and the two communities, 'classical' biochemists and molecular biologists, begin to merge.

One increasingly exploited strategy is that of site-specific mutagenesis. By this is meant an assortment of chemical procedures that selectively modify individual nucleotides, resulting in single amino acid substitutions in the primary translation product. This selective change in protein sequence permits detailed investigation of protein structure/function relationships. Such techniques have been developed by many biochemists including Dr. H.G. Khorana, who also used this technology in the work described in a highlight below. Bacteriorhodopsin is probably the best studied membrane protein, and, since the sequence, gene, and function are known, it lends itself to site-specific mutagenesis as a tool to dissect protein conformation and the associated proton transport. The studies by Dr. John Gerlt, also highlighted below, show the value of this approach. This investigator has been able to deduce at least some aspects of the reaction mechanism by combining mutagenesis, with structural information gained from X-ray diffraction.

Another major technical tool, whose use has been spreading to all areas of modern biomedical research, is that of monoclonal antibodies. These antibodies have the advantage of great specificity, a specificity so high that the antibodies usually react only to a small epitope on the appropriate antigen. Thus, it is possible to dissect an antigen sequence, or domain, by raising a large

number of monoclonal antibodies. The monoclonal antibody technology is not without its drawbacks. In addition to the time necessary to raise and screen for useful lymphocyte clones, the antibodies themselves generally have lower binding affinities than standard rabbit or goat polyclonals, and the fact that they are generally monovalent limits their use. Nevertheless, they have proven immensely useful. Several highlights below rely on the use of monoclonal antibodies for their exciting findings. Although in the past few years we have included a few highlights using monoclonal antibodies, this is the first year we have several, and they reveal some of the breadth of utility of these antibodies. For example, Dr. J. Glorioso has used monoclonals to dissect the glycoprotein coat of herpes virus into its several domains. These domains are of great interest because some are conserved between different types of herpes virus, and others are type specific. Such dissection of domains should lead to understanding of the mechanism of viral entry into target cells, as well as the development of vaccines to the envelope glycoproteins. A second example is the use of monoclonals by Dr. A.F. Horwitz. To find a membrane protein of a given function, in this case cell-cell and cell-substrate attachment, monoclonals are screened for interference with that function. Here, the CSAT strain was found to disrupt cell attachment. Standard biochemical techniques are then used to find the target antigen, which did indeed turn out to be an attachment protein. Finally, Dr. N.B. Gilula used monoclonal antibodies against a known antigen of identified function, in this instance gap junction protein, to selectively disrupt the function of that antigen. Since selective inhibitors are not always known, and their selectivity always open to question, the specificity of monoclonals holds the promise of providing the perfect inhibitor. Although in practice the selection and use of such inhibitors turns out to be very difficult, using these antibodies, Dr. Gilula has demonstrated some important and interesting consequences of disruption of intercellular communication in embryos.

Another technical advance which has been supported by the CMBD Program and used by grantees is the development of calcium-sensitive dyes, most notable by Dr. R. Tsien. Over the past several years the annual report has included discussion on the role of intracellular calcium as a regulator of cell processes, and a putative second messenger. Grantees have looked at the control of gap junction permeability by calcium, as well as that of exocytosis and cell motility. It has been clear that intracellular calcium is very important, but its activity is very low, perhaps below 10^{-7} M in many cells. This makes the development of probes which are selective and sensitive in this range all the more important but difficult. Dr. Tsien has also synthesized agents which transport calcium, and several other grantees have synthesized and studied calcium ionophores. Often the mode of action of intracellular calcium is via calmodulin and related proteins. NIGMS supports some excellent research into calmodulin biochemistry; the highlight on Dr. D.M. Watterson discusses a particularly promising example.

This year we include two highlights on cell cycles. In the OBJECTIVES section above, we describe the wide range of research supported by CMBD. Although some of the theoretical studies on model systems and diffraction analyses of proteins now are supported by the Biophysics and Physiological Sciences Program, the CMBD program still stretches from the atomic basis of enzyme catalysis, to cell development, growth, and cell-cell interactions. In many ways, research at the highest level of organization, the cell or tissue, is most difficult because the systems are so complicated and the variables hard to know or control. Yet, cell biology has made remarkable strides in the last decade, and these two

highlights provide some insights into the hypotheses. The research of Dr. C. Helmstetter, although into 'cell-cycles', actually concerns bacterial replication. As we know more and more about bacterial physiology, and have been able to manipulate the organisms, the bacteria have proven to be useful models for processes that ordinarily are identified with higher organisms. One example we have provided this year is the use of Myxobacteria as models for developmental processes. Dr. A.D. Kaiser has proven the usefulness of these bacteria as models, and other grantees have utilized B. subtilis sporulation, or Caulobacter stalk growth, as developmental models. Dr. C. Helmstetter has investigated cell replication in E. coli, and has provided insights into the sequence of events which may also be found in eukaryotes. Dr. A. Pardee, on the other hand, has used eukaryotic cells, although in tissue culture, to ask questions about the control of the cell cycle. He, and many other CMBD grantees, some of whom have been cited in this highlight, have found tissue culture cells to be well-suited to the study of cell cycles. It is possible to control growth in these cells, synchronize entire cultures, and manipulate the stages of the cycle. As can be seen in the highlight, under these growth conditions it is possible to investigate in detail the molecular biological basis of cell function and the genetic basis of metabolic processes and cell growth can be elucidated.

NIGMS is the basic research institute, and supports many lines of study of interest to more than one categorical institute. In general, we would expect a period of several years between NIGMS-sponsored research and the beginning of application of this knowledge to actual disease processes. In some cases, the techniques and approaches developed under NIGMS auspices may be utilized in research projects supported by the categorical institutes, which in turn lead to understanding, treatment and prevention of disease. Given this, it might seem surprising how clearly research supported by this institute relates to normal and abnormal human function. However, the gap between the most basic research and human function is never that great, and the CMBD highlights included below provide several further examples of that point.

The research by Dr. S. Wakil, for example, reveals how detailed examination of the enzymes in a pathway in bacteria can shed light on the function of a human enzyme. In this case, the enzyme is in the pathway of synthesis of fatty acids, a process of immediate concern to millions of Americans. Dr. Bernard Trumpower has shown the advantage of using a bacterial model to study a process, mitochondrial respiration, that is so complicated that it is hard to study in mammalian tissues. Dr. J. Glorioso's research is of obvious potential usefulness. Although his work employs the virus as a model, herpes simplex and its relatives are important human viruses, whose mode of transmission, mechanism of cell penetration, and antigenicity are significant questions in medical research. Dr. Khorana's work, although directed primarily to bacteriorhodopsin, is now being extended to vertebrate rhodopsin. This protein, and the closely related iodopsins, are the human visual pigments. Although essential for vision, and intensively studied, surprisingly little is known about important features of rhodopsin, and Dr. Khorana promises to add appreciably to our understanding of these proteins. These research efforts demonstrate the utility of bacteria as models.

Calcium plays such an important role in many human cell processes, and calmodulin is an essential intermediate in so many of these, that its human

significance needs little added. One point to be made is that calmodulin is very similar to calcium-sensitive muscle proteins, and several CMBD grantees are currently studying the regulation of calcium fluxes in muscle cells. These fluxes control the actual contractile force generation. The similarities between cellular motility and muscle motility are many, including the role of calcium-sensitive protein in modulating both muscle contraction and ciliary movement, as detailed in the work by Dr. Whitman. The ATPase containing force-generating protein in cilia, dynein, presents an interesting comparison with muscle myosin. Drs. Gilula and Horwitz, whose work has broad potential import, have already shown that the proteins they are studying play an essential role in normal vertebrate development. Dr. D. Goodenough has been studying the Zona Occludens. This subcellular structure is found in many human tissues, but he has been finding out a great deal about its role in the liver, where it is essential to proper epithelial transport.

The highlight on anion transport cites many workers, and reveals how NIGMS support of basic function, in this case ion transport, leads to the understanding of essential human physiology, here gas exchange in the lungs and blood. Although it is not surprising that the use of human erythrocytes as models will lead to the understanding of the function and disease of human blood, what is unexpected is the renewed interest in anion transport in general. NIGMS has supported some key studies on the genetics of cystic fibrosis, but generally we have not supported much research into the physiological defect. However, research supported by other institutes has increasingly pointed to the primary defect as being in anion transport. Work sponsored by us as well as by others, has revealed many of the common features of anion transport in the diverse transporters found in nature. This may mean that much of the research on the erythrocyte anion transporter may be applicable to the chloride transport deficiency found in cystic fibrosis.

CMBD supports an appreciable amount of research into cell-cell recognition and adhesion. The work of Dr. S.D. Rosen, highlighted below, demonstrates how important such recognition can be. In this case, cell recognition is essential to the systemic immune response. Taking an approach initially developed with slime molds, this investigator is now investigating mammalian immune response.

Lastly, Dr. Pardee's research demonstrates how cultured cells can be used to ask questions concerning growth control and cancer. Since cancer is a disease of growth control, research into growth can lead to understanding of cancer. In addition, Dr. Pardee, and some of his colleagues, also supported by NIGMS, have led in the study of oncogenes and their role in the control of the cell cycle. This may lead to our understanding of the function of proto-oncogenes in normal cells, as well as the essential steps in cell transformation.

As can be seen from these descriptions, CMBD supports a wide range of projects, which develop and utilize state-of-the-art techniques to further our understanding of basic biomedical processes. Such understanding can be of great, and even immediate, significance.

"The T4-bacterial Gene Coding for Thymidylate Synthase"
RO1 GM 26387-06 (Maley, F.), New York State Department of Health, Albany

Thymidylate synthase is an important enzyme in DNA synthesis in that it catalyzes the methylation of deoxyuridylate to thymidylate, an essential precursor

for DNA biosynthesis. The enzyme has been a target for chemotherapeutic drugs since the finding that this enzyme is the primary site of action of 5-fluorouracil, an analog of the nucleotide substrate. Unfortunately, because this drug impairs the synthase from tumor and normal tissue to the same degree, the ability to obtain a selective response in favor of the host is rather limited. However, a great degree of selectivity was observed with analogs of the second substrate of thymidylate synthase, 5, 10-methylenetetrahydrofolate (mtf) when the synthases from T4-bacteriophage and its bacterial host, Escherichia coli, were compared. Thus, it was observed that the phage enzyme could be completely inhibited by an analog of mtf without affecting the host synthase. To understand the molecular basis for this effect, Dr. Frank Maley and his collaborators undertook a structural comparison of these enzymes. Dr. Maley had previously established the amino acid sequence of the Lactobacillus casei synthase by classical protein techniques; however, he determined the sequence of the bacteriophage and E. coli enzymes by examining both the proteins and their genes. To aid in this procedure the gene for the bacterial enzyme and that for the phage enzyme (td) were isolated and placed in a plasmid vector capable of expressing these genes and their protein products in large quantities. Although no great surprises were found on comparing the sequences of the E. coli and L. casei enzymes, a 51 amino acid deletion in the E. coli synthase was found, which when omitted from a comparison of the two enzymes revealed a 62 percent homology between them. This homology increased to 82 percent in the active site region.

In contrast to the E. coli gene, which encompassed a 792-base pair open reading frame coding for 264 amino acids, the gene, as isolated, for the phage enzyme contained 2,094 base pairs while encoding for only 286 amino acids. On the basis of three nucleotides per amino acid, only 858 nucleotides should have been required to provide sufficient information to code for the T4-phage synthase. This anomaly was resolved by the finding that the gene was interrupted by an intron of 1,017 nucleotides two-thirds of the way through the open reading frame for the coding sequence of the phage gene. The intron contained a termination codon at its 5'-end; at its 3'-end the intron possessed a start codon, followed by an open frame for the remaining 102 amino acids encoded by the gene (see the schematic below). The sequence of the phage enzyme deduced from its gene was confirmed by Dr. Maley using protein sequence data. This finding was the first reported for the presence of an intron in the structural gene of a non-eukaryote.

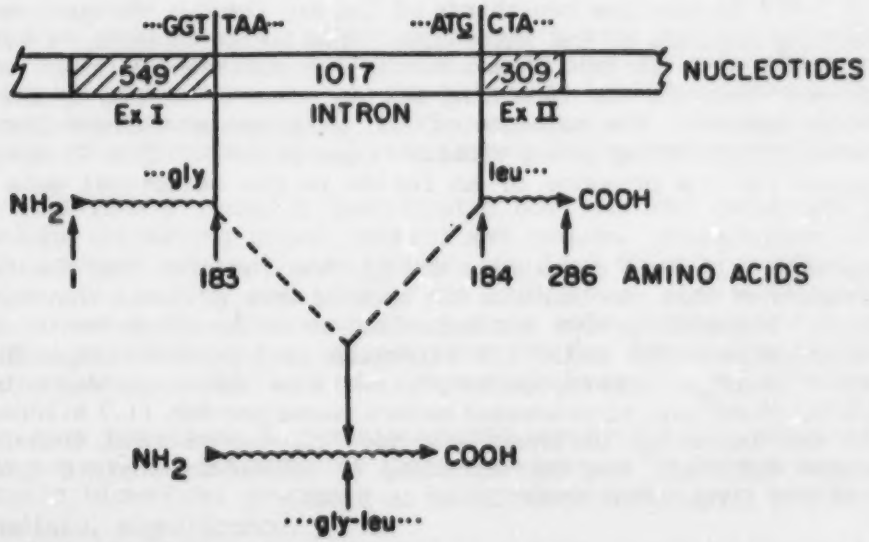
In vivo studies with a plasmid containing the td gene revealed that during the course of expression of this gene various RNA species were produced that hybridized with labeled oligomeric probes corresponding to different segments of the gene. Thus, a premessage RNA about 2.7 kilobases in length corresponding to the primary RNA transcript was recognized, as well as those corresponding to the intron (1.0 kilobase) and to processed mature messenger RNA (1.7 kilobases). Coincident with the formation of these RNA species, two protein translation products were also observed, one corresponding to mature thymidylate synthase and the other to the first exon, about 25 kD in size.

Similar, but even more dramatic findings were obtained by Dr. Maley's laboratory using an in vitro coupled transcription-translation system. They found not only that the two protein products described above could be immunoprecipitated with antibody to T4-phage thymidylate synthase, but also that active enzyme synthesis could be detected. In keeping with what was anticipated, enzyme appearance followed that of mature message, which in turn followed the appearance

of pre-message. Characterization of the RNA products produced *in vitro* with the SP6-RNA polymerase system revealed a complex mixture of RNAs containing pre-message, message, and both of the exons, which are normally spliced to form mature synthase message. More recently it was shown that incubation of the isolated pre-message (2.7 kilobases) in a test tube, under various conditions, results in excision of the 1.0 kilobase intron consequent with the union of the two exons to form mature thymidylate synthase message. The excised intron is present in both circular and linear structures. Although similar results have been shown by Cech (GM 28039) with *Tetrahymena* ribosomal RNA, it is the first time that this route for mature message formation has been shown outside of a eukaryotic system.

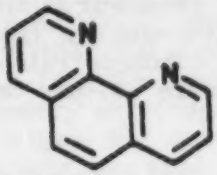
Aside from these striking results, the mechanism of which is still under investigation, another significant finding was obtained. The original goal, that of explaining the difference in inhibition of the T4-phage and *E. coli* synthases, was approached when Dr. Maley's group identified an arginine in the folate binding site of *E. coli* thymidylate synthase, but found a lysine in the same site in the T4-phage synthase. The *L. casei* synthase, which was also subject to inhibition by folate analogs, also contained a lysine in its folate binding site. It would therefore appear that the placement of an arginine at the binding site could contribute to the ability of the *E. coli* enzyme to escape inhibition by the folate analogs.

Thus, while the original intent of this investigation was to conduct a phylogenetic comparison of various thymidylate synthases, it has revealed some fascinating, albeit unanticipated, findings concerning the structure of a viral gene.



"Clarifications of DNA Polymerase Chemistry"
ROI GM 21199-16 (Sigman, D.), University of California, Los Angeles

The chelating agent, ortho-phenanthroline (OP), has played an important part in studying the role of transition metal ions in biological catalysis.



ortho-phenanthroline (OP)

Inhibition of enzymatic activity by this chelating agent has been taken as an indication that a metal ion might play a central role either in the catalytic mechanism or the structure of the enzyme. Beginning in the early 70s, many laboratories reported that OP inhibited a variety of DNA and RNA polymerases; this observation and other data have led to the widely-held idea that these polymerases are dependent on zinc ions for their mechanistic and/or structural integrity. It has now become apparent, however, that OP inhibition of at least one polymerase, the well-studied E. coli DNA polymerase I (pol I), is far more complex than originally thought. The nature of this inhibition, which has been elucidated by Dr. David Sigman, is the subject of this highlight.

Dr. Sigman has found that the inhibition arises not from OP itself but rather from a complex it forms with adventitious copper ions, carried into the assay mixture as a contaminant in the buffers. This OP-copper ion complex, in the presence of a reducing agent and the oxygen present in solution, generates peroxide which, under these conditions, causes DNA chain cleavage. This cleavage reaction produces, among other species, DNA fragments phosphorylated on the 3' terminus, and it is these materials which are actually inhibitory. These fragments, of course, readily bind to the enzyme because they have the general structural features required by the pol I of a template-primer, but the presence of the phosphate on the normally-reactive 3'-hydroxyl group precludes their utilization as a substrate.

Dr. Sigman is investigating the details of this reaction and has found that the reactive species (1) reacts preferentially with double-stranded DNA that is in the B conformation (Z-DNA, interestingly, is unreactive), (2) has no effect in the presence of agents known to intercalate, and (3) is capable of chain cleavage between any pair of adjacent nucleotide residues, although local conformations within B-DNA can render particular regions unreactive. The former two observations suggest that the catalytic species, which is actually the tetrahedral 2:1 OP-cuprous ion complex, must itself bind to DNA in order to facilitate the scission reaction. Dr. Sigman has proposed a plausible mechanism, involving hydroxyl radicals produced from peroxide in the presence of the complex, that account for the observed properties of this reaction.

This reaction of the OP-cuprous ion complex and peroxide with double-stranded DNA is of interest from several vantage points. First, this work defines a previously unrecognized mode of inhibition of polymerases by OP. Secondly, hydroxyl radicals and other oxygen species appear responsible for many DNA modifications, including chain scission; thus, the further examination of this reaction may reveal some principles of these important DNA-damaging processes.

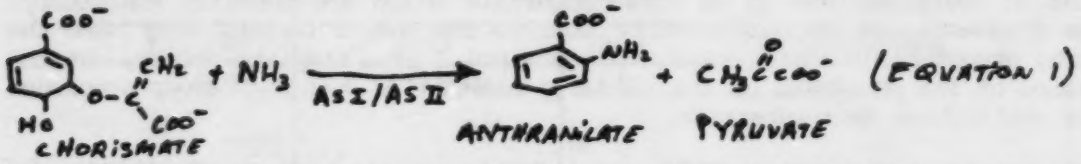
Third, since the nuclease activity of the OP-cuprous ion complex appears to be dependent on the three-dimensional structure of the nucleic acid, the reagent may be useful in delineating local conformational variations in DNA structure.

It should also be pointed out that a number of laboratories, including Dr. Sigman's, have, in view of these findings, carefully re-examined the Zn-ion dependency of DNA pol I and found that neither the polymerase nor the 3'→ 5' exonuclease activity is truly Zn-ion dependent. These findings raise obvious questions for the many other polymerases thought to be Zn-ion dependent.

"Work on Glutamine Amidotransferases"
R01 GM 24658-15 (Zalkin, H.), Purdue University
R01 GM 09738-23 (Yanofsky, C.), Stanford University

Glutamine amidotransferases are a widely-distributed class of enzymes that utilize the side chain $-CNH_2$ group of glutamine as a nitrogen source in the biosynthesis of amino acids, purine and pyrimidine nucleotides and certain coenzymes. These enzymes, which play a significant role in the metabolism of assimilated nitrogen by both eukaryotes and prokoryates, have been studied for many years by Dr. Howard Zalkin and his colleagues at Purdue University. While Dr. Zalkin has investigated the structural and functional properties of a number of glutamine amidotransferases, this summary will highlight only one of the enzymes, anthranilate synthetase.

Anthranilate synthetase (AS) is perhaps the best characterized of the glutamine amidotransferases. The enzyme, which catalyzes the first committed step in tryptophan biosynthesis, is an oligomer of nonidentical subunits, referred to as AS I and II. The AS I subunit alone can catalyze the ammonia-dependent synthesis of anthranilate (equation 1) via a multi-step process.



The AS II subunit binds glutamine and imparts to the AS I/AS II complex the ability to use glutamine as an amino nitrogen source. The rate of anthranilate production via the AS I/AS II complex and glutamine is higher than via AS I and ammonia. AS is typical of glutamine amidotransferases in that it is also divided into two functional units. One unit catalyzes hydrolysis of the glutamine amide to make ammonia immediately available to the second unit. This latter unit takes ammonia obtained either from the first unit or from solvent and adds it to some metabolic intermediate (such as chorismate: see equation 1). In AS these functional units are associated with separate proteins but in other glutamine amidotransferases they exist as separate domains of a single protein.

Chemical modification studies of the AS I/AS II complex have strongly implicated an AS II active site cysteine as being essential for the glutamine amidotransferase activity, and, as expected for an active site residue, this cysteine is conserved in all microbial AS II sequences. Dr. Zalkin, some years ago, proposed a chemical mechanism involving the cysteine that accounts for the capacity of the AS II subunit to free the amide nitrogen from glutamine and transfer it to AS I for incorporation into anthranilate.

Recently, Dr. Zalkin and his co-workers have used site-directed mutagenesis to examine further the role of the cysteine in glutamine utilization. Working with AS II subunit from the bacteria Serratia marcescens, they replaced the implicated cysteine (at sequence position 84) with a glycine residue and found that the replacement abolished the glutamine-dependent anthranilate synthetase activity of the mutant AS I/AS II complex but not its ammonia-dependent activity. While the most likely explanation of these results is that the catalysis is dependent on the cysteinyl residue, one can question whether the functional loss results from the replacement of the putative catalytic residue or results from a secondary structural alteration. However, recent structural comparison of the wild-type and mutant enzyme in Dr. Zalkin's lab has indicated very strong similarities; thus, the loss of activity can be attributed to the absence of a catalytic component, the cysteinyl residue.

The AS II subunit from S. marcescens is a small (M.W.= 20 kD) monofunctional protein, as it is for a number of other bacteria. However, in some other prokaryotes the AS II subunit is part of a larger bifunctional protein; in these prokaryotes the other portion of the protein catalyzes the second step of the tryptophan pathway. The AS II subunit structure of two eukaryotic enzymes has now been examined by Dr. Charles Yanofsky (GM-09738) at Stanford University with collaboration from Dr. Zalkin. They have determined that the AS II subunit from these eukaryotes is also part of a multi-functional enzyme. For example, these workers have isolated and sequenced a yeast gene that codes for a protein composed of an N-terminal sequence homologous to a known AS II subunit, an eleven residue linker peptide, and a C-terminal sequence homologous to a protein with indole glycerolphosphate synthetase (IGPS) activity. (This activity is responsible for one of the last few steps in tryptophan biosynthesis.) While the enzyme hasn't been isolated and characterized, this gene sequencing work virtually establishes the existence in yeast of a bifunctional protein responsible for both AS II and IGPS activity. Thus, it is likely that this organism has the capacity to modulate coordinately the level of these two enzymatic activities, which are both needed in tryptophan biosynthesis. This conclusion is strengthened by the finding that in the 5'-region flanking the coding sequence there is a nucleotide sequence that could function in the regulation of gene transcription.

Finally, Dr. Zalkin has noted from his sequence work on AS and other glutamine amidotransferases that a homology exists between many of these enzymes in that portion of the protein that hydrolyses the glutamine; this homology is consistent with a common evolutionary origin of these structures. It has further been suggested that an early organism "recruited" the prototypic glutamine hydrolyzing protein for use by the ammonia-requiring enzymes, thereby acquiring a survival advantage from the obvious added metabolic flexibility.

"Progress on Understanding the Enzyme Fatty Acid Synthetase"
RO1 GM 19091-14 (Wakil, S.), Baylor College of Medicine

The seven enzymes which catalyze the biosynthesis of the fatty acid palmitate from acetyl-CoA and malonyl-CoA in E. coli are known to be associated in vivo; they are, however, dissociated into discreet individual enzymes upon cell disruption. The protein components (seven enzymes plus the acyl carrier protein) and the individual reactions of palmitate biosynthesis are summarized below:

(1) Acetyl transacylase (AT)

$$CH_3COS\text{-}CoA + ACP\text{-}SH \rightleftharpoons CH_3COS\text{-}ACP + CoA\text{-}SH.$$

(2) Malonyl transacylase (MT) ACP = Acyl carrier protein

$$\begin{matrix} COOH \\ | \\ CH_2COS\text{-}CoA \end{matrix} + ACP\text{-}SH \rightleftharpoons \begin{matrix} COOH \\ | \\ CH_2COS\text{-}ACP \end{matrix} + CoA\text{-}SH.$$

(3) β-Ketoacyl-ACP synthetase (KS)

$$CH_3COS\text{-}ACP + Enz\text{-}SH \rightleftharpoons CH_3COS\text{-}Enz + ACP\text{-}SH.$$

$$CH_3COS\text{-}Enz + \begin{matrix} COOH \\ | \\ CH_2COS\text{-}ACP \end{matrix} \longrightarrow CH_3COCH_2COS\text{-}ACP + CO_2 + Enz\text{-}SH.$$

(4) β-Ketoacyl-ACP reductase (KR)

$$CH_3COCH_2COS\text{-}ACP + NADPH + H^+ \rightleftharpoons D\text{-}CH_3CHOHCH_2COS\text{-}ACP + NADP^+$$

(5) β-Hydroxyacyl-ACP dehydratase (DH)

$$CH_3CHOHCH_2COS\text{-}ACP \rightleftharpoons trans\text{-}CH_3CH{=}CHCOS\text{-}ACP + H_2O.$$

(6) Enoyl-ACP reductase (ER)

$$CH_3CH{=}CHCOS\text{-}ACP + NADPH + H^+ \longrightarrow CH_3CH_2CH_2COS\text{-}ACP + NADP^+$$

The product of reaction six is utilized as a substrate for repetitions of the elongation sequence (reactions 3-6) until the palmitate chain is fully elaborated to give $CH_3(CH_2)_{14}COS\text{-}ACP$

(7) Thioesterase (TE)

$$CH_3(CH_2)_{14}COS\text{-}ACP + H_2O \longrightarrow CH_3(CH_2)_{14}COOH + ACP\text{-}SH.$$

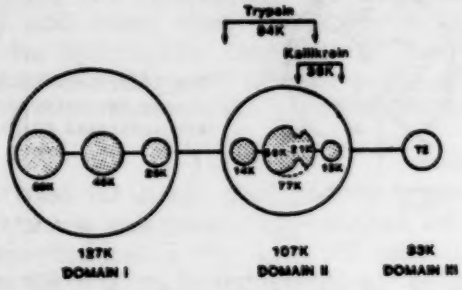
While these same reactions are utilized for this pathway in vertebrates, the individual catalytic activities that facilitate these reactions are not associated with discreet proteins. They are, rather, the properties of a massive dimeric protein. In its native form, this protein, fatty acid synthetase (FAS), catalyzes all the individual steps in palmitate biosynthesis. Recently, a model that incorporates most of the structural and functional properties of FAS was proposed by Dr. Salih Wakil of the Baylor College of Medicine. This model, which was worked out over the last decade with the support of the NIGMS, is the subject of this highlight.

By the mid-1970's there was general agreement, based on the work of a number of investigators including Dr. Wakil, that FAS was composed of two subunits which were identical in all respects; each subunit had the same molecular weight, shape, charge, and molar content of the prosthetic group, 4'-phosphopantetheine. Interestingly, the dissociated subunits retained six of the seven enzymatic activities required for fatty acid synthesis but lacked the β-ketoacyl synthetase (KS) activity—the activity required for carbon-carbon bond formation.

Working out of this background, Dr. Wakil is carrying out two related studies—one, to determine the structural properties of the subunits, and the other, to ascribe functional characteristics, i.e., enzymatic activities, to the structural regions defined in the first study. Structural studies of multifunctional enzymes often reveal that they are arranged as a series of globular domains that contain the sites of catalytic or regulatory activity; these domains are

connected by polypeptide bridges which are often more exposed and therefore susceptible to proteolytic cleavage. This susceptibility sometimes makes it possible to fragment a protein into separable domains, and such is the case for fatty acid synthetase. Indeed, Dr. Wakil has found that there are three major domains, referred to as I, II, and III. The latter one of these is a 33 kilo-Dalton(kD) protein specifically removed from the FAS subunit by α- chymotrypsin leaving a 230 kD core protein. Functionally, the Domain III protein was easy to characterize--it was a thioesterase and presumably responsible for removing the fatty acid end product from the enzyme in vivo. The core protein as a dimer, retained all the repetitive activities involved in the step-wise assembly of fatty acyl chains but lacked the ability to release the product. Since the intact synthetase had a blocked N-terminus, and the Domain III protein has a free N-terminus, it was concluded that Domain III was located at the carboxyl end of the subunit.

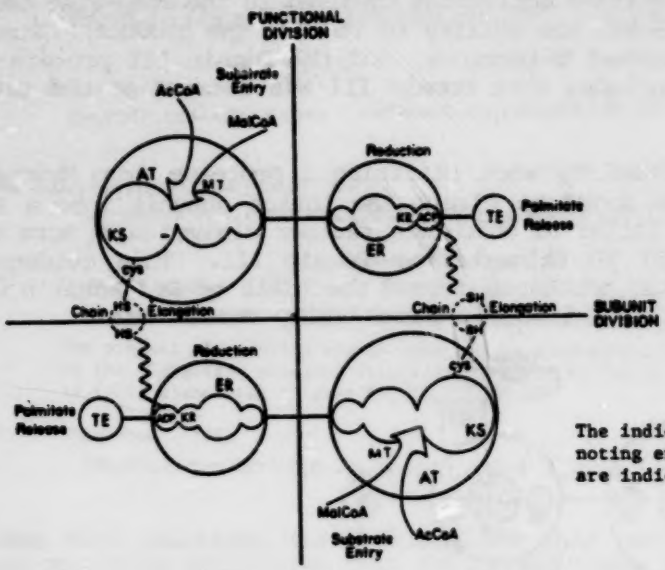
The other domains were defined by work utilizing a protease from Myxobacter. This proteolytic enzyme was found to cleave the intact subunit into a 127 kD and a 138 kD fragment, the latter of which was further cleaved much more slowly to a 107 kD protein and 33 kD thioesterase--Domain III. This evidence and similar cleavages using other proteases formed the basis of Dr. Wakil's map of the structural regions of FAS, which are schematically shown below:



As the figure indicates, the larger domains may be further subdivided and mapped by the action of individual proteases into a number of distinct regions: Domain I into sub-domains of 60, 45, and 23 kD; Domain II into sub-domains of 14, 56, 21, and 15 kD. This proteolytic map served as a reference for the determination of the loci of the functional activities of the synthetase.

Starting with this structural model, as embodied in the fragmentation map, this group has sought to locate the functional centers of the enzyme. Although functionality in a multifunctional protein is not necessarily associated with specific proteolytically-generated fragments, it certainly can be, as illustrated by the association of thioesterase activity with Domain III. Dr. Wakil's laboratory has thus far been able to make assignments (with differing degrees of confidence) of six of the seven catalytic activities of FAS. They have also located the site to which the growing fatty acid chain is anchored during synthesis. These assignments depended upon the known properties of the active sites of the individual activities. For example, the growing fatty acid chain is known to be anchored to the enzyme through the thiol of the phosphopantetheine (PPT) prosthetic group; thus, when FAS labelled with ^{14}C-PPT was proteolytically fragmented, it was observed that the label was exclusively found in the 15 kD (carboxy terminal) portion of Domain II. This evidence strongly indicates that

this portion of the large vertebrate enzyme serves as the analog of the acyl carrier protein of the coli system. In another case, the α-ketoacyl reductase activity was found to be expressed fully by two different, albeit overlapping, proteins, the 94 kD and 36 kD fragments of Domain II indicated in the figure. This observation and other confirming evidence localized the α-ketoacyl reductase activity to the 21 kD stretch that constituted the overlapping sequences of the 94kD and 36 kD fragments. Using studies such as these, Dr. Wakil has been able to ascribe (see figure 2) all of the functions of the FAS system except the dehydratase activity to delimited areas of the protein.



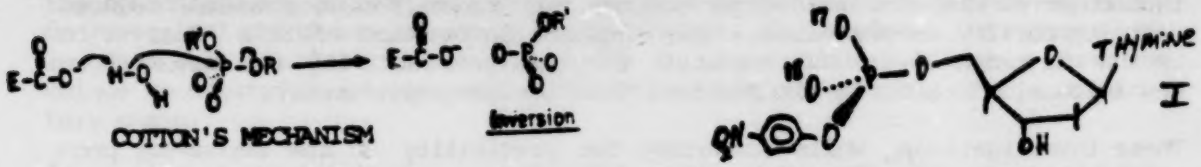
The indicators (ER,KR,etc.) noting enzymatic activity are indicated on page one.

As the figure indicates, Dr. Wakil is proposing that the enzyme has two functional palmitate synthesizing units, each of which consists of Domain I from one subunit and Domain II and III of the other subunit. The basis for this head-to-tail association lies largely with the findings that α-ketoacyl synthetase activity is dependent upon the association of the two subunits and the established proximity of a critical -SH group of a Domain I fragment and the -SH of the PPT moiety located in Domain II. Work on these maps of FAS is still underway in Dr. Wakil's laboratory and will no doubt produce further insight into the structural and functional properties of this important and very complex enzyme.

"Studies of Staphylococcal Nuclease"
RO1 GM 34573-02 (Gerlt, J.), University of Maryland, College Park

The enzymes that catalytically establish or cleave phosphate ester bonds are immensely important to cellular function in that they play crucial roles in the biosynthesis and degradation of nucleic acids, in the capture and release of energy for cellular function, and in the transport of materials across cellular and organelle membranes. As part of a program to understand the principles that govern the function of this class of enzymes, Dr. John Gerlt has been studying the nuclease secreted by Staphylococcus aureus. This enzyme is a Ca^{+2}-dependent phosphohydrolase that catalyzes the hydrolysis of

single stranded DNA and RNA. This enxyme is particularly amenable to detailed mechanistic studies because it is quite well characterized structurally. Dr. C. Anfinsen and co-workers established the amino acid sequence in 1971, and Dr. R. Cotton reported a 1.5 A X-ray structure of the enzyme, complexed to an inhibitor, in 1979. A chemical mechanism for the nuclease was proposed by Dr. Cotton based on the crystal structure. The center feature of Dr. Cotton's mechanism, shown in the figure below,



was the action of the $-COO^-$ group of Glu-43 to facilitate the attack of water on the phosphate ester. This mechanism, if operative, demands that an inversion of stereochemistry occur if substrates that are chiral at phosphorus are employed. The mechanism also demands that oxygen from solvent water end up in the phosphate product. Another plausible mechanism, still consistent with the X-ray data, would predict different results for each of these experiments. Dr. Gerlt has examined the hydrolysis by the nuclease of compound I (shown above on the right) and found that inversion at phosphorus indeed occurs. He has also examined the hydrolytic reaction in H_2O^{18} and found that the enzyme does catalyze the incorporation of O^{18} into the phosphate product. These results clearly strengthen the credibility of Cotton's mechanism for the nuclease.

Dr. Gerlt has continued to study nuclease by adopting site-directed mutagenesis techniques to produce convenient quantities of "mutant" enzymes which were used to evaluate experimentally the role of various residues that may participate in catalysis and/or substrate binding. The gene for the nuclease, which was cloned and sequenced by Dr. David Shortle (GM 34171), was inserted into an expression-secretion plasmid. In this plasmid, which was constructed in collaboration with Dr. Masayori Inouye (GM 19043), the nuclease gene was directly fused to the leader sequence of the outer membrane A protein produced by E. coli, and the resulting construct placed under the control of a lac promoter. The plasmid directed the synthesis and secretion of large amounts of the nuclease into the periplasmic space of the E. coli host. Dr. Gerlt was then able to substitute into the plasmid-borne gene any of a number of synthetic oligonucleotides to yield, after expression, the desired mutant enzymes.

The first mutant prepared by Dr. Gerlt's laboratory contained an aspartate at sequence position 43. The Asp-43 mutant hydrolyzed DNA 200-times slower, and displayed a binding affinity for DNA that was five times weaker than the parent enzyme. The simplest explanation for these observations is that the α-COO^- of Glu-43 is directly involved in catalysis as indicated by all prior data, and that the β-COO^- of the Asp-43 can serve as a poor but still chemically effective base in the catalytic process. Refinements of this explanation may be possible when the X-ray crystallography of the Asp-43 mutant nuclease is complete.

Dr. Cotton's X-ray structure of the nuclease suggested that the phenolic hydroxyl group of tyrosine-85 interacts with substrates via the 3'-phosphate

group of nucleotide substrates; kinetically, the interaction of substrates with the nuclease also implicated the presence of an acidic functional group having a pK_a about 9. In addition, chemical modification studies performed by Dr. Anfinsen and his colleagues implicated this tyrosine residue as being important in catalysis. Dr. Gerlt and his co-workers have constructed the phenylalanini-85 mutant enzyme, and the initial kinetic characterization of this mutant enzyme shows it to have a V_{max} value essentially identical to that observed for the wild type enzyme; the enzyme has a somewhat reduced affinity for DNA as substrate. The simplest explanation of this observation is that tryosine-85 is not essential for catalytic activity (in contrast to the impression created by the chemical modification experiments).

These investigations, while enhancing the credibility of the mechanism proposed by Dr. Cotton, illustrate the value and interdependency of the numerous techniques used to define enzyme mechanism. The role of site-directed mutagenesis among these techniques is to provide mutant enzymes in quantity and with speed and high specificity; conclusions, of course, drawn from the mutant comparisons of the two.

"The Respiratory Chain of Paracoccus"
RO1 GM 20379-12 (Trumpower, B.), Dartmouth Medical School

Respiration, in biochemical terms, is the process by which energy-rich "fuel" molecules are oxidized to provide cellular energy, in the form of ATP. In the oxidation process, electrons are removed from the fuel and passed ultimately to some acceptor, commonly oxygen. In eukaryotes, respiration takes place in the mitochondria via a very complex membrane-dependent system of many proteins. The complexity of this electron-transporting oxidative phosphorylation system has made the respiratory process difficult to elucidate, and investigators have sought simpler related systems in hopes of gaining insight into eukaryotic mitochrondrial respiration. One such simple respiratory system is found in the plasma membrane of Paracoccus denitrificans, a bacteria which normally grows in soil. In this niche, oxygen content can be quite variable. If adequate oxygen is unavailable Paracoccus will use various nitrogen compounds in its respiration process as electron acceptors--converting these to nitrogen gas in a process known, not surprisingly, as denitrification. If oxygen is present, Paracoccus elaborates a respiratory chain which appears to be functionally identical to that of eukaryotic mitochondria. Under these aerobic conditions this bacteria can oxidize carbon-containing compounds identical to those used by mitochondria, and reduce molecular oxygen to water. The presence of a "mitochondria like" respiratory chain in other bacteria is relatively rare and, for this reason, Paracoccus has attracted the interest of researchers investigating the molecular basis of respiration and energy transduction in eukaryotic cells. Among the leading investigators in this area is Dr. Bernard Trumpower, who has undertaken the isolation and characterization of the Paracoccus respiratory chain and its components.

In 1983, Dr. Trumpower and his coworkers discovered that the functionally-intact respiratory chain of this bacteria could be isolated in fully active form from the rest of the plasma membrane. Functionally, this aggregate of lipids and proteins, termed a quinol oxidase complex, oxidizes ubiquinol to ubiquinone and transfers the electrons acquired in the process to molecular oxygen. The proteins of this aggregate that mediate this electron transport consist of the bc1 complex, a membrane-bound cytochrome c (c-552), and the

cytochrome aa_3 complex. The order of electron transfer in this aggregate is shown in the figure. The really striking property of this quinol oxidase

ubiquinol ⟶ bc, complex ⟶ cyt C_{552} ⟶ cyt aa_3 complex —
ELECTRON TRANSPORT CHAIN OF PARACOCCUS (cyt = cytochrome)

complex is that it consists of only six protein subunits. In mitochrondria, the cytochrome bc_1 and aa_3 complexes, which, together with cytochrome c, carry out ubiquinol oxidation, consist of approximately 20 proteins. This comparison suggests that the "extra" proteins found in mitochondria may have a function unique to eukaryotic cells, possibly in biogenesis and assembly of the respiratory chain.

In continuing their studies on the Paracoccus chain, Dr. Trumpower and his colleagues have purified the cytochrome bc_1 complex, which is the constituent of the quinol oxidase complex that transfers electrons from ubiquinol to cytochrome c-552 (see figure above). The Paracoccus bc_1 complex consists of only three proteins, in comparison to nine in mitochondria. It has the smallest number of protein constituents, and one of the highest turnover numbers of any bc_1 complex isolated from any species to date.

One excellent example of the advantages which this bacterial respiratory chain complex may offer for investigations into the mechanism of respiration comes from recent experiments on cytochrome b conducted in Dr. Trumpower's laboratory. A barrier to understanding how this particular cytochrome functions in respiration has been the inability to purify this protein, a problem which has confounded numerous investigators in this area for many year. However, workers in this laboratory have found that cytochrome b could be purified from the Paracoccus bc_1 complex in a single step.

This purification is based on the phenomenon of temperature dependent phase separation of the detergent Triton X-114. When the three subunit bc_1 complex is mixed with this detergent at 0°C, it forms a uniformly dispersed mixture. When the temperature is raised to 25°, the detergent undergoes a microscopic phase separation to form a cloudy emulsion. When this emulsion is clarified by low speed centrifugation, the detergent is recovered as a separate phase, and is found to contain pure cytochrome b. The second, aqueous phase, contains the remaining two subunits, cytochrome c_1 and the iron-sulfur protein of the bc_1 complex. This separation is quantitative, can be performed in minutes, and can be reversed by lowering the temperature.

It is clear from these early results with Paracoccus that this quinol oxidase complex and the three subunit cytochrome bc_1 complex offer exciting opportunities to study a "simplified" respiratory chain. As demonstrated above, the bc_1 complex appears especially amenable to resolution into separate components and reconstitution. The ability to reconstitute bc_1 complex using mutant b protein produced through recombinant techniques is a particularly promising avenue for elucidating the functional characteristics of that protein.

"Genetics, Biochemistry and Biology of Herpes Virus Glycoproteins"
RO1 GM 34534-05 (Glorioso, J.), University of Michigan Medical School

The Herpes viruses are responsible for a wide range of diseases associated with

lesions of the skin, mucosal membranes, and central nervous system. In this country, the HSV-1 viruses are primarily linked to facial lesions (i.e., the common cold sore), ocular keratitis (associated with blindness), and acute necrotizing encephalitis, whereas the HSV-2 virus is largely responsible for genital Herpes. The related Herpes zoster virus is responsible for childhood chicken pox, and shingles in the adult. Thus, studies directed towards the biological, biochemical, and genetic properties of the Herpes viruses have broad implications for the diagnosis and treatment of the diseases for which they are the causative agent.

Dr. Glorioso and colleagues at the University of Michigan, Ann Arbor, have been studying the major glycoproteins of the HSV-1 viral envelope, which are encoded by the viral genome and expressed as membrane components on virally infected cells. These studies are aimed at understanding the genetics of the Herpes viruses, and dissecting the role of their glycoprotein molecules in viral infection and host resistance. The experimental approach has been to generate a panel of monoclonal antibodies which recognize type specific epitopes of the HSV-1 envelope. These monoclonal antibodies have been used to characterize both the antigenic structure of the viral proteins and to select mutants lacking these epitopes, which differ in their biological properties.

Using this approach, four HSV-1 encoded glycoproteins have been identified and designated gB, gC, gD and gE. The gB glycoprotein has a molecular weight of 123 kD with a peptide backbone of 900 amino acids. The molecule has an unusually long, hydrophobic transmembrane domain which may play a role in penetration of the virus into the cell. There is some evidence that this transmembrane component also forms part of an ion channel. Viral mutants which lack the gB determinant adhere to cells but do not penetrate the membrane, and thus cannot infect the cell and multiply.

The gC glycoprotein is of equal interest and importance, in that it appears to play a key role in the host's immune response to HSV-1 infection. The gC molecule is even larger than gB, with a molecular weight of 130 kD, due to a high degree of glycosylation of the polypeptide chain. This molecule, which is encoded by a single gene, is structurally interesting in that it has both a type specific divergent domain, as well as a highly conserved domain shared by HSV-1, HSV-2, and Herpes zoster.

The gC glycoprotein appears to play a significant role in the host response to HSV-1 infection. One of the interesting aspects of Herpes infections is that the virus often lies dormant for long periods of time within the nervous system, providing a reservoir for recurrent infection. This may be just one of several ways by which the Herpes viruses escape the immune response. The gC molecule bears a receptor for the third component of complement, C_3b, and thus probably plays a role in the complement mediated immunologic lysis of virally infected cells. Furthermore, this molecule expresses antigenic determinants which serve as targets for cell mediated natural killer, T-cell, and antibody dependent lysis of virally infected cells.

Mutants have been isolated which are gC negative by virtue of the fact that the gC molecule lacks its cytoplasmic tail and therefore is secreted, rather than integrated, into the cell membrane. Generation of these mutants could provide an important mechanism for the escape of virally infected cells from immunological surveillance.

The other two glycoproteins that have been characterized, gD and gE, are smaller (58 and 83 kD respectively) and much less is known of their biological roles. However, with the use of monoclonal antibodies directed to specific epitopes, it has been shown that gE bears an Fc receptor for immunoglobulin and that gD plays a role in adhesion of HSV-1 to cells. Continuing studies of the HSV-1 glycoproteins, and the monoclonal antibodies directed against them, should lead to a more complete picture of the genetic, and biological parameters of the Herpes virus family. These studies already have proven useful in the diagnosis of Herpes infections by epitope and type specific antibodies.

"Synthetic and Biological Study of Nucleic Acids"
RO1 GM 28289-05 (Khorana, H.G.), Massachusetts Institute of Technology

An important class of proteins contains retinal, vitamin A aldehyde, as a chromophore. This group includes eukaryotic visual pigments, most notably mammalian rhodopsin. At least three distinct retinal-containing pigments are found in the halophilic archaebacterium, Halobacterium halobium. The best-studied of these is bacteriorhodopsin (BR), which contains the all trans isomer of retinal, attached to a single integral membrane protein, bacterio-opsin (BO). Illumination isomerizes the retinal to 13-cis. This is the only light-dependent step, and is followed by a series of dark reactions. The complex photocycle involves a series of conformational changes in the retinal, which contains a system of conjugate double bonds. The energy is transmitted to the protein moiety and results in the pumping of a proton by BR across the membrane. This light-driven protein pumping establishes a pH gradient which can provide energy for other cellular processes. Under appropriate conditions H. halobium produces membranes with dense patches of BR, in a nearly crystalline array. This BR is of great interest as a model system for a variety of processes. First, it shares several characteristics with human visual pigments. Second, it is a good model for the study of biochemical photochemistry. Third, it is of interest as a membrane transporter and especially as a proton pumper. Proton pumping is an important process in bioenegetic reactions in prokaryotes, plants, and animals. Finally, its abundance and array in bacterial membranes allows preparation of large quantities, and study of its spatial organization in and near the bilayer. It is, perhaps, the best studied intrinsic membrane protein.

Several NIGMS grantees have made important, even seminal, contributions to the study of BR. Dr. W. Stoeckenius, of University of California, San Francisco, is one of the founders of the field and, with his colleagues, continues to make major findings in all aspects of BR biochemistry and biophysics, and H. halobium physiology. Dr. J. Lanyi, at University of California, Irvine, has also been a long term contributor to the field. He has recently pioneered the study of a second retinal pigment in H. halobium, named halorhodopsin. This pigment appears to serve as a light-driven chloride pump. Recently, a third pigment has been studied by Dr. John Spudich, of Albert Einstein Medical School, among others; this pigment seems to play a central role in photomechanical transduction, mediating the bacterium's phototropic response. Dr. Nigel Unwin, of Stanford University, has used the crystalline membrane arrays of BR to produce a 3 dimensional map of its tertiary structure, through the technique of electron diffraction, which he helped perfect. Many other first-class investigators have used diffraction and other physical and chemical techniques to study BR structure. In fact, the entire BR field, as well as that of the two other recently-discovered pigments, has attracted first-class investigators.

Among those is Dr. H.G. Khorana, of MIT, who established an international reputation in the area of nucleotide chemistry.

Dr. Khorana has turned his expertise in organic chemistry, biochemistry, and molecular biology to the study of BR. One surprising finding is the ability of BR to be denatured and totally reconstituted. For example, BO can be boiled in SDS or treated with trifluoroacetic acid. Upon appropriate treatment, the protein refolds accurately and quantitatively to the native structure. Retinal can then be rebound, and the native chromophore regenerated quantitatively. This BR can be reconstituted into liposomes to generate a fully efficient proton pump. BR is also unique among intrinsic membrane proteins in the ability to recombine chymotryptic fragments to form the native protein. Combining these techniques, Dr. Khorana has been able to alter a fragment, (for example by deuteration), for physical analysis of the structure. Another major finding has been the identification of the site of retinal attachment as lysine 216. This has been followed up by photoactivatable crosslinking experiments which have mapped out the orientation of retinal within the protein.

Simultaneously, the Khorana lab has been studying the molecular biology of BO and its genetic control. They have shown that it is synthesized as a precursor with a 13 aminoacid leader sequence. The synthesis, membrane integration, and terminal clipping are all cotranslational. Using the known protein sequence, a cDNA probe was isolated, and the gene characterized. It contains no introns, but does code for an extra aspartate residue at the carboxyl end. As the first archaebacterial gene sequenced, many of its features have received a great deal of interest. One problem with experimentation with an archaebacterial gene is the lack of an efficient expression vector system. The Khorana lab has been able to achieve a low level of expression of the BO gene in E. coli. Despite the low yield it has been possible to prepare milligram quantities of BO from E. coli. The lab is now preparing a large number of mutants through several mechanisms of site-specific mutagenesis. Simultaneously, an effort is being made to develop a homologous expression system. The ultimate questions to be asked are the relationship of sequence to structure and function. Since we know more about BR than any other membrane protein, the ability to manipulate the sequence should provide very sophisticated answers to the structure-function question.

One of the advantages of studying BR and the other H. halobium retinal pigments is that they share many properties with vertebrate rhodopsin. Dr. Khorana has recently extended his research into the molecular biology of rhodopsin. He has isolated and sequenced cDNA clones containing the coding sequence for 90 percent of the amino acid sequence of cattle rhodopsin, and has incorporated it into an expression vector. Efficient expression in E. coli was achieved. The laboratory has also been studying the structure and structure-activity relationships for rhodopsin. Dr. Khorana's current plans are to totally synthesize the rhodopsin gene. His lab is probably the leader in gene synthesis. This is to be followed by specific nucleotide alterations, to essentially produce true designer proteins. This promises to be a major step in our understanding of rhodopsin, one of the most important and interesting vertebrate proteins.

"Biochemistry and Radiation Effect During the Cell Cycle"
RO1 GM 26429-20 (Helmstetter, C.), Roswell Park Memorial Institute

When a cell undergoes mitosis and divides into two daughter cells, it is essen-

tial that each of these cells contains equal and identical amounts of genetic material. Thus the DNA contained within the chromosomes of the cell about to undergo mitosis must replicate in an orderly way prior to cell division. Cellular multiplication is controlled to a significant extent by the mechanism which determines the frequency of initiation of chromosome replication. Research on the biochemical mechanism of initiation of chromosome replication therefore is essential not only for characterizing the fundamental processes of cell and tissue growth and development, it is also of importance for understanding and treating diseases which involve aberrant cellular multiplication, such as infections and cancer.

Dr. Helmstetter, who has been supported by NIGMS for the past 20 years, has devoted much of his research effort to studying the control mechanisms involved in initiation of chromosome replication, and the role of this event in the overall control of the cellular cycle of growth and division.

In Dr. Helmstetter's most recent work, significant advances have been made in understanding the factors which determine the frequency of initiation of chromosome replication, using the simple bacterium, Escherichia coli, as a model system. This organism contains a single chromosome which initiates DNA replication from a specific site called oriC. Under a given set of growth conditions, DNA replication initiates from oriC at a precisely determined time in the cellular division cycle.

Dr. Helmstetter's latest studies on the control of initiation from the oriC replication complex have been facilitated by the analysis of minichromosomes which Dr. Helmstetter and others have constructed. These minichromosomes contain only 500-1500 base pairs, and include a resident copy of oriC; as many as 50% of these minichromosomes can replicate in cell free extracts in approximately 20 minutes. The small size and accessibility to in vitro study of minichromosomes of Escherichia coli will facilitate experiments which are likely to yield a great deal of information on the precise biochemical events which take place at the replication origin at the time of initiation.

In the current phase of this project, minichromosomes are being used to address several important questions in control of replication:

1) Role of genes adjacent to oriC--since the complete DNA sequence of the minichromosomes under study is known and can be modified, it is possible to explore the involvement of genes adjacent to oriC in initiation of replication. For example, it has been found that genes encoding 16 kD and 17 kD peptides to the right of oriC are needed for the maintenance of minichromosomes; cells with minichromosomes lacking this region cycle more slowly.

2) Characterization of cell-cycle dependent proteins which interact with the oriC complex - using minichromosomes in which genes from temperature sensitive mutants had been inserted, Dr. Helmstetter has shown that, following shifts to intermediate temperatures, synchronous bursts of initiation occurred in the minichromosomes. These bursts required protein synthesis. The proteins have been partially characterized and their rates of synthesis, stability, and conservation during the cell cycle are now under study.

The results so far suggest that an initiation complex (oriC and associated genes and proteins) can be used only once, probably because it is degraded. In

addition, there is evidence that initiation/replication depends upon gene activation, the product of which could be an anti-repressor for the long postulated repressor of initiation. Thus, these studies using minichromosomes are providing information on a positive control system for replication mediated by one or more of the proteins required for initiation.

The long-term goal of this project is to identify the specific macromolecular component of the cell which is the rate limiting determinant of the timing of initiation of replication from oriC. It is anticipated that the details of the macromolecular events at initiation, and the elements which control the frequency of these events, will be determined in large part through the study of the replication properties of minichromosomes.

"Control of Multicellular Development in Bacteria"
RO1 GM 23441-09 (Kaiser, A.D.), Stanford University

The development of a multicellular organism raises many fundamental questions. Among them: What regulatory mechanisms underlie cellular differentiation? How is the time sequence of developmental events controlled? How do individual cells coordinate their behavior so that each cell plays its proper role at the proper time?

In answer to the last question there is evidence that cell behavior is coordinated by the exchange of signals between cells. For example, in vertebrate embryos the formation of the lens involves interactions between different cell types. Developmental signals are also found in the immune system, where soluble factors transmit growth and differentiation signals between various types of lymphoid cells. Any new information on intercellular chemical signals and their control of developmental pathways would be of great value.

It would seem that the last organisms to be useful in the study of development and intercellular signaling would be unicellular ones. Moreover, among unicellular organisms, bacteria would be least likely to be of interest. Recently, however, it has become clear that a group of gram-negative bacteria, the Myxobacteria, undergo a primitive differentiation and a true multicellular development in its life cycle based on chemical signaling between cells. Among the first investigators to study this development, and still a leader in the field, is Dr. M. Dworkin, of the University of Minnesota (GM 19957). Another laboratory investigating this very interesting system has been that of Dr. A.D. Kaiser, of Stanford University (GM 23441).

Dr. Kaiser has set out to answer the following overall questions: What are the intercellular chemical signals used to coordinate the development of Myxococcus xanthus? What are the genetic and regulatory relationships among the genes that control Myxococcus colony or life cycle development; is its development organized by a dependent regulatory pathway? How do the intercellular signals control development and how do they fit within the entire deveopmental regulation of fruiting bodies? The advantage of a bacterial model is that it is very accessible to genetic, molecular, biological, and biochemical manipulation and analysis.

Dr. Kaiser has isolated four classes of developmental mutants, SpoA, B, C, and D, and analyzed them with respect to the stage of development affected. One of

the series, SpoC, has been cloned in E. coli. SpoC negative cells can be rescued totally, producing a wild-type number of spores which appear to be normal by all physical and biochemical criteria. Rescue is achieved by addition of four murein components: N-acetylglucosamin, N- Acetylmuramic acid, diaminopimelic acid, and D-alanine. Dr. Kaiser has also developed culture methods which produce fruiting bodies in submerged culture. Thus, it is now possible to add components to liquid cultures to complement Spo mutants. The rescue of SpoC mutants in submerged culture shows the value of this system. During normal development at least half the cells lyse, releasing peptidoglycan and enzymes that degrade it into its components. It seems possible that these chemical components are normal intercellular developmental signals.

Dr. Kaiser has also made interesting discoveries about the genetic control of fruiting body development in the colonies. In genetic investigations, this researcher has shown that a transposable element, Tn5, can transpose to many different sites in the Myxococcus genome. It thus can provide a selectable marker for strain construction, localized mutagenesis, chromosome mapping, and cloning. The development of sophisticated and powerful genetic techniques for Myxococcus is recent. There are already many interesting biochemical observations and more classical genetic findings which will be well worth following up using molecular biological techniques. One example is the finding that starvation for amino acids not only initiates development but also elicits accumulation of guanosine polyphosphates. The possible causal relationship is of great interest.

Dr. Kaiser has developed a first class bacterial model for development. He has improved culture conditions, isolated developmental mutants, and now developed vectors and the techniques for cloning genes in E. coli plasmids. This simple developmental system might be the first one to become fully understood in terms of genetic and molecular mechanisms.

"Regulation of Cell Processes by Calcium"
RO1 GM 30861-04 (Watterson, D.), Vanderbilt University
RO1 GM 31004-04,(Tsien, R.), University of California, Berkeley

Calcium is essential for the normal functioning of most living cells. Many physiological processes, such as clotting, secretion, muscle contraction, and wound healing, are dependent upon the presence of calcium. At the molecular level, numerous enzymes, hormones, and other biochemical mediators require calcium in order to be biologically active. Thus the molecular basis of calcium action, and the role it plays in the regulation of cells is an important biological issue.

Several research groups supported by the CMBD program are devoting their activities to the elucidation of the mechanism of calcium action.

The laboratory of Dr. D. M. Watterson, at Vanderbilt University, Tennessee, has been investigating a class of intra-cellular proteins, called calmodulins, which reversibly bind and release calcium for use in a multitude of calcium-dependent cellular processes. Calmodulin is ubiquitous, found in almost all eukaryotic organisms, and its structure and composition is remarkably conserved throughout the phylogenetic scale, suggesting its fundamental role in basic life processes. The actual molecular mechanism of how calmodulin functions to regulate calcium-

dependent processes is the subject of Dr. Watterson's research. In an attempt to solve this problem, he has recently constructed a gene which can be used for the synthesis of the calmodulin protein molecule. It will now be possible, by selective modification of this gene, to alter small regions of calmodulin and determine how these changes influence the binding and release of calcium. Insertion of the gene into cells will enable detailed studies of how this molecule regulates calcium-dependent cellular processes.

In contrast to the studies of Dr. Watterson, which address the role of calcium bound to proteins, Dr. Tsien is investigating the role of free intracellular calcium in the regulation of cell growth and proliferation. Dr. Tsien has developed new techniques which offer unprecedented sensitivity and spatial resolution for the detection of free cytosolic calcium. This has been achieved by designing fluorescent calcium-binding dyes which can be quantitatively assessed using the fluorescence-activated flow cytometer. These dyes are a vast improvement over Quin 2, which was previously used by investigators in this field, and will be invaluable to a broad range of calcium studies.

Dr. Tsien plans to apply this new methodology to several important biological questions, including changes in free calcium following fertilization in the dividing sea urchin zygote; activation of resting lymphocytes by mitogenic and non-mitogenic lectins, lymphokines, antibodies and antigens; and stimulation of fibroblasts by growth factors. These studies should result in new information on the role of free calcium in mitotic and non-mitotic cell events.

"Molecular Mechanism of Flagellar Motility"
RO1 GM 30626-05 (Whitman, G.), Worcester Foundation for Experimental Biology

Flagella and cilia are locomotory organelles composed of a constant structural arrangement consisting of nine double peripheral microtubules and two single central microtubules. These organelles, which project and move in waves from the surface of certain cells, have become classic models for motility, in great part because of their microtubular composition. Understanding the mechanism of flagellar movement is significant in its own right because of the important function flagella and cilia play in man, and because the microtubules, which play a key role in their activity, are also important in many other cell functions. Microtubules have been shown to be essential for cell movement, play a major role in cell division, and assist in the transport of biochemicals into and out of the cells. Large numbers of cilia are found on cells which line the respiratory tract where they function to screen out dust and other foreign particles. Both cilia and flagella play an important role in human reproduction: the coordinated beating of cilia in the oviduct produces a current which draws the egg into the uterus, while the rapidly moving tail of the spermatozoa, a flagellum, contacts the egg and results in fertilization.

The dynein arms generate the forces that are the basis for motility in all eukaryotic cilia and flagella, including those of man. It has now been well-established by investigators in the field that the mechanochemical basis for flagellar movements is the ATP-driven sliding of adjacent doublet microtubules. Dynein-ATPase containing arms on the outer doublets provide the motive force for microtubule-sliding in flagella. It is to be emphasized that dynein arms generate the forces that are the basis for motility in all eukaryotic cilia and flagella.

Recent studies, from the laboratories of Dr. Whitman and others, have revealed that the dynein arms are complex structures containing multiple subunits and at least three sites of ATP-binding and hydrolysis. For this research, the algal flagellate Chlamydomonas reinhardtii is used. The long-term goals of this project are to determine the molecular structure of the outer arm of the Chlamydomonas flagellum. Specifically, his laboratory is providing interesting data regarding the alpha and beta subunits of 18S dynein found in the arms. They are determining the relationship of the subunits to the whole 18S particle and subfractionating the beta subunits to learn more about their individual chains and their relationships to one another. New monoclonal antibodies specific for the dynein chains have been developed and will be used to determine the locations of the dynein chains, to identify the structural domains involved in binding of dynein to microtubules, to investigate the functions of the three outer arm dynein ATPases, and to clarify the relationships between the dynein of different species.

Dr. Whitman and others have also been studying the location and function of the tightly bound calmodulin in the axoneme of Chlamydomonas, in order to determine if calmodulin also binds to the axoneme in a CA++-dependent manner. The axonemal proteins involved in both calcium-independent and calcium-dependent binding of calmodulin will be identified. The possibility that flagellar waveform is controlled by a phosphorylation-dephosphorylation mechanism will be investigated.

Dr. Whitman has clearly established himself as one of the leading investigators in the area of flagellar motility and dynein biochemistry. His laboratory is impressive in its utilization of an array of state-of-the-art techniques in protein biochemistry, ultrastructure, and immunological and genetic approaches in providing answers regarding the molecular and structural organization of flagellar outer dynein arms and the functions of the individual subunits and polypeptide chains of its dynein ATPase.

These studies will not only answer basic motility questions but will provide a basis for understanding such important processes as sperm maturation and capacitation, which involve changes in the functioning of the arms, and chromosome movement in dividing cells, which also appear to involve dynein-like ATPases.

Why is determining and understanding the structure and function of cilia and flagella important? Ciliary defects occur in humans in various forms, such as male sterility which results from immobile sperm. Depending on the form of the hereditary disease, the sperm may lack structural parts, such as both dynein arms, either the outer or inner arms, the radial spoke heads, or the inner sheath together with one or both of the central pair of microtubules. These same defects also occur in the respiratory cilia of individuals who commonly have long histories of respiratory tract disease, i.e., recurrent bronchitis and chronic sinusitis. These conditions have been shown to result from immobile respiratory cilia which are unable to do their task of clearing mucus from the lungs and sinuses. This "immobile cilia syndrome" appears to be an inherited condition and it has been observed that approximately 50 percent of individuals with it also have a rare condition known as "situs invertus", in which there is a reversal of the normal symmetry of the body, i.e., location of organs as in the position of the intestines, heart, lungs, etc. This entire complex of abnormalities is known as Kartenger's syndrome. Basic research studies of cilia and flagella, and in the case of the "immobile cilia syndrome" and Kartenger's syndrome, specific studies of the protein dynein, will provide

answers, explaining not only the cause of the defect, but will also hopefully provide answers leading to the restoration of normal function.

Every new bit of information about the molecular make-up of the cell's organelles is one step more toward the development of better drugs for the treatment of diseases and ways of perhaps even preventing some of the diseases which plague us now. Once we know what controls the cell, we can begin to understand how to repair it or treat it when it malfunctions.

"The Role of Gap Junctions in Development"
ROI GM 32230-04 (Gilula, N.) Baylor University

In its last report, the CMBD Program highlighted the major advances in the understanding of gap junctions between cells. Such direct cell-to-cell communication (cell coupling) is a feature of virtually all animal tissues. The coupling is known to provide a low conductance electrical pathway between neurons, muscle cells, and among many other cell types. Although many adult cells are not coupled, they are derived from coupled embryonic cells. There appears to be a special role of coupling in development. It has been proposed that the coordination of cells during embryogenesis is mediated by something passing between the cells through the gap junctions. The junctions not only pass electric currents (that is, ions), but also larger molecules, including metabolites and messengers.

As noted in the previous report, many workers are addressing the question of the mechanism of control of the normal closing and opening of the gap junction channels. For example, Dr. Camillo Peracchia of the University of Rochester (GM 20113) has provided much evidence on the role of calcium in modulating the coupling. A slight increase in intracellular calcium can uncouple the cells. Dr. P.N.T. Unwin, of Stanford University (GM 28668, GM 27764 and GM 30387) has used a combined X-ray and electron diffraction approach to determine the structure of gap junctions. He has also shown that a slight increase in calcium can alter the gap structure such that the apparent channel appears to be occluded.

Although cells have developed a sensitive and elegant means of controlling intercellular coupling, it is difficult for investigators to investigate this process. A long term uncoupling would require, for example, maintenance of elevated intracellular calcium levels. However, the cell has also evolved mechanisms for keeping intracellular calcium low. In addition, calcium can mediate many other intracellular processes, such as calmodulin-dependent phosphorylation. Therefore, a more specific and permanent blocker of gap junction channels would be a very useful tool. In addition, despite the channel seen in many, but not all, structural studies of gap proteins, the connection between cell coupling and the presence of gap junctions is still inferential. A specific blocker of coupling which also specifically interacts with gap structures would add powerful new evidence to the presumed, but unproven, identity of gap junctions as the actual site of intercellular molecular transport.

Dr. Gilula appears to have succeeded in making just such a selective probe/blocker. He has made several polyclonal antibodies against the major, 27 kD Dalton protein from rat liver gap junction. It is notable that polyclonal antibodies were far superior. Monoclonal antibodies are extremely specific to small epitopes on the antigen. Unfortunately, it has proven very difficult to obtain monoclonals directed against the functional portion of the 27 kD protein.

In any case, Dr. Gilula has shown using electron microscopic immunolocalization that the polycolnal antibodies selected can bind to the cytoplasmic surface of intact rat liver gap junctions, using electron microscopic immunolocalization. These antibodies also were shown to cross react with a 54 kD protein in Xenopus (frog) eggs and embryos. Xenopus embryo cells from the 8 or 16 cell stages appear to exhibit normal coupling even after injection of control preimmune rabbit serum. This was demonstrated by the transfer of the intracellular marker dye, Lucifer yellow, between cells. Antibody injection blocks the spread of the dye from the treated cell. Measurement of electrical coupling, a more sensitive gauge of intercellular connection, revealed that antibody injection reduced or totally eliminated all coupling.

The evidence is certainly consistent with the concept of the gap junctions being the actual site of intercellular molecular transfer. The antibodies seem to be a specific probe of the gap coupling sites. Most interesting, the injection of antibodies generates a high proportion of development abnormalities which were related to the developmental fate of the injected cell. There was even a clear pattern to the nature of the abnormalities. Injection of antibody generated a high proportion of tadpoles with a characteristic set of abnormalities related to the developmental fate of the injected cell. The most frequently occurring defect was one of asymmetric development. For example, the right eye would fail to form and there would be a gross underdevelopment of the brain on the right side. These experiments are a significant step, if only a beginning, towards our eventual understanding of the role of intercellular communication and exchange in development.

"The Erythrocyte Anion Transport"
RO1 GM 22432-11 (Chan, S.), California Institute of Technology

The most heavily used ion transport protein in typical vertebrates is the anion transporter of the erythrocytes. Even the abundant sodium/potassium ATPase of kidney and brain, or the proton-ATPase of mitochondria transport fewer ions. The anion transporter catalyzes the chloride-bicarbonate exchange in the erythrocyte. In the systemic capillaries, CO_2 enters the capillary lumen by diffusion; it then diffuses into the red cell and is very rapidly converted to the HCO_3 and H^+ by carbonic anhydrase; the high intracellular bicarbonate drives a bicarbonate-chloride exchange. The net effect is to convert gaseous CO_2 in the systemic capilaries into plasma bicarbonate. The entire process is reversed in the lungs. In lungs or systemic tissues, capillary transit time is less than half a second. Thus, the elements of this exchange, the anion transporter and intracellular carbonic anhydrase, must catalyze a rapid process and be sufficiently abundant to handle the respiratory load. Therefore, it is not surprising that there are many copies of the anion transporter.

Fractionation of red cell ghosts, which contain just a membrane and membrane-associated matrix proteins and lipids, produces a large band on SDS acrilamide gels, band 3, which can be shown to be the anion transporter. This peptide is a 95,000 dalton integral membrane protein that constitutes about 25 percent of the total red cell membrane protein. Two NIGMS grantees, O. Frohlich and R. Gunn of Emory University, have estimated, on the basis of kinetic data, that there are about one million transporters per cell. Freeze fracture by another NIGMS grantee, D. Branton of Harvard University, shows intramembranous particles widely accepted as being mainly band 3. Estimates of the numbers of such

particles is usually around 400,000 per cell. There is, therefore, some reason to believe that the particles are aggregates.

The anion transporter differs from most, if not all, other ion transporters in several of its kinetic characteristics. First, although chloride and bicarbonate are transported most readily and rapidly, many other anions can also permeate via band 3. This includes several monovalents and some divalents as well. In fact, another NIGMS grantee, Dr. A. K. Solomon has recently presented substantial evidence that water and urea cross erythrocytes through the same 'chloride channels'. The chloride and bicarbonate tranport is unusual in that it consists almost entirely of an obligatory one-for-one exchange. Measurement of ^{36}Cl-Cl exchange shows that this is about ten thousand times more rapid than would be predicted from chloride conductance mechanisms. This, together with other evidence, makes it very unlikely that the anion transporter is an ionic channel. Although the wide range of transported anions, rapid transport, and high turnover rate per site would be consistent with movement through a large (several Angstrom) anion channel, this is clearly not the case. The identification of membrane fraction band 3 as the anion exchanger has allowed molecular dissection of transport function although, unfortunately, the total sequence is not yet known. The N-terminal consists of a large hydrophilic domain facing the cytoplasmic surface. It can be cleaved enzymatically, leaving a hydrophobic peptide which comprises about 55 percent of the original band 3 protein, yet still retains near normal transport function. Since band 3 is the most abundant integral membrane protein, it is not surprising that it serves as an attachment site for linker proteins which bind to the cytoskeletal matrix, thus providing the erythrocyte ghost with its unusual rigidity. NIGMS grantee, Dr. V. Bennett, of Johns Hopkins University, has shown that one major such linker, ankyrin, attaches to the cytoplasmic domain of band 3. The C-terminus is heterogeneously glycosylated, although the reason for this variability is unknown. Many workers have applied covalent reagents, or peptidases, to either membrane face to reveal the pattern of transmembrane crossing, as well as to produce fragments for structure-activity studies. Using this technique, one NIGMS grantee, Dr. M. Jennings, of the University of Iowa, has demonstrated that a small, 17,000 dalton, chymotryptic fragment spans the lipid bilayer more than once. His evidence, and that of others, indicates that the band 3 monomer crosses the bilayer several times, possibly seven, thus resembling the best-studied intrinsic membrane protein, bacteriorhodopsin, which contains 7 alpha-helical segments, each of which crosses the entire bilayer.

The unusual kinetics of the anion transporter has been of great interest. A large family of potent inhibitors has been synthesized, all of which are derivatives of stilbene disulfonate. These compounds, usually known by their initials, such as SITS, DIDS, DNDS, and BIDS, do not penetrate the membrane, but inhibit the transport when applied in the extracellular solution. Although a few derivatives appear to be competitive, most react covalently, and irreversibly inhibit anion transport. Dr. A.K. Solomon has been one of the investigators most intensively studying the nature of stilbene disulfonate derivative inhibition. Most recently, Dr. S. Chan, of California Institute of Technology, has also been using these compounds, and other probes, to investigate the transport mechanisms. He has developed a $^{35/37}Cl$ NMR assay. This affords a direct spectroscopic probe of the actual anion site, and he and his colleagues have used it to provide evidence for "alternate site" models of transport, such as the "ping-pong" mechanism.

Dr. Chan has observed ^{35}Cl NMR line broadening in the presence of membranes containing band 3 protein. The evidence that this is binding to the transporter includes the following: 1) the sites are inhibited competitively by DNDS; 2) The sites have affinities for DNDS and Cl^- that are quantitatively similar to the known affinities for the pair; 3) The sites have relative affinities for Cl^-, HCO_3^-, F^- and I^- that are quantitatively similar to the known relative affinities of these ions for band 3 transport. The NMR probe has been used to investigate the transport cycle. If a noncompetitive inhibitor, such as pNBS, is applied to resealed ghosts containing a low internal concentration of readily tranportable anion, only a fraction of the transporter sites are blocked for subsequent transport [after pNBS is washed out and appropriate intracellular anions are restored]. This implies that, in the absence of a transportable intracellular anion, some of the band 3 proteins remain with their binding site facing the inside surface, and thereby inaccessible to the covalent binding of the stilbene derivatives which only bind to the transporter when the anion-binding site is facing outward. Under conditions where inward facing sites transport acions out, all sites are blocked by pNSA. Dr. Chan has noted that this, and his other kinetic data, are consistent with the "ping-pong" model of Gunn and Frohlich, in which a single transport site alternates between the inward- and outward-facing states and can only change states when occupied by bound anion. From NMR data, Chan has been able to set the limits of the rates for chloride binding and dissociation at both faces. The binding and dissociation in each direction exceed 100,000 events/sec/site. This is much faster than the known turnover rate of about 430 at 0°C. Assuming that the unidirectional, half turnover rates do not differ by more than one hundred fold, then it can be shown that the translocation of the chloride-transport site complex across the bilayer is the rate-limiting step. Dr. Chan and his colleagues have further combined Cl-NMR with inhibitor studies to provide evidence of the active chloride binding sites. Covalent, arginine-specific reagents, phenylglyoxal and 1,2-cyclohexanedione, eliminate the transport site NMR linebroadening. Further experiments show this linebroadening to be pH-sensitive, with a pKa of about 11.1. Thus, there is strong evidence that an arginine residue, accessible to these reagents, provides the positive charge in at least one conformation of the transport site. Dr. Chan now plans to follow up these experiments by binding labeled phenylglyoxal to the band 3 arginine, digestion with papain, and location of the labeled residue. Further work will address not only the binding site, but perhaps the most interesting question of all, how the structure and composition of this site changes during translocation. The anion transporter is so interesting and important to human function, its study will be a major focus of his and other excellent laboratories for several years.

"Biochemistry of Cell Attachment"
RO1 GM 28932-05 (Goodenough, D.), Harvard University
RO1 GM 23244-10 (Horwitz, A.), University of Pennsylvania

Very few cells are freely moving and unattached. Rather, most cells adhere to other cells and/or the extracellular matrix. The attachment is often extensive, and provides the tissue with its characteristic stiffness and rigidity. In some cases attachments may serve not only as mechanical links, but also as permeability barriers to the free movement of materials through the intercellular cleft. Some attachments are extremely labile, forming and disassembling during development, perhaps governing the morphogenesis of the cell itself or the en-

tire tissue. The control of this lability is of great interest in the study of development, and metastasis as well.

Epithelia which separate physiological compartments act as permeability barriers to the transepithelial passage of ions and molecules. This passage, called the paracellular pathway, is sealed by a gasket-like intercellular junction, the Zona Occludens (ZO). Electron microscopy resolved the ZO into a series of punctate fusions between adjacent cells' plasma membranes, which branched and anastomosed in the plane of the membranes, generating a web-like structure, very characteristically displayed with freeze-fracture electron microscopy. Using electron dense tracer, it can be shown that the ZO represents a barrier to large molecules transiting the paracellular space. It is presumed that this is also a major component of the ionic and osmotic barrier which is essential to normal epithelial function. In the past decade evidence has accumulated that the ZO barrier and structure can be abolished and reformed, perhaps subject to control by calcium activity. Additionally there is substantial evidence that several cytoskeletal components, such a α-actinin and vinculin, are associated with the ZO, perhaps anchoring in or close to the intercellular junction.

Dr. Daniel Goodenough, of Harvard University, and his colleagues, have been studying the ZO for several years. He has found that standard mouse liver preparations contain intact bile canaliculi and hepatocyte "bile front" membranes; and the bile canaliculi contain ZO's along with other junctions such as desmosomes, gap junctions, zonulae adherentes, as well as lateral plasma membranes. Treatment with the detergent, DOC, in combination with the calcium ligand, EGTA, results in the solubilization of most of the membranes, leaving the junctional complexes as long ribbons, visible by negative staining. Further extraction with the detergent, n-lauroyl sarcosine, reduces the residue but leaves the junctional complex ribbons and solubilizes or denatures the ZO fibrils. SDS-PAGE analysis of the DOC-insoluble pellet reveals relatively few bands. Only a few of these are solubilized by sarcosine, and these are the leading candidates for components of the ZO fibrils. Although it has been possible to raise polyclonal antibodies to some of these proteins, this lab has been unsuccessful at obtaining antibody to the fibrils. A specfic rabbit antibody to the junctional complex has been raised, and fluorescence microscopy has confirmed its specificity for part of the complex. However, colloidal gold localization in the EM has shown the antibody to be associated with some other specific junctional element. However, any component of the junctional complex is of interest. Although mouse liver junctions have proven far more amenable to purification and fractionation than those from other species tried, they are, not surprisingly, not effective antigens for raising mouse monoclonal antibodies. Therefore, Goodenough and his colleagues are switching to the fractionation of beef liver, with the hopes that large quantities of junctional complexes can be obtained, the components of which will be effective at raising mouse antibodies.

A slightly different approach has been taken by Dr. A. Horwitz, of the University of Pennsylvania, and his collaborators. They have been investigating a monoclonal antibody, CSAT, which was found to specifically inhibit the adhesion of chick myoblasts to the cell substratum. Immunohistochemical techniques localized the CSAT antigen at the periphery of the cell along stress fibers, a staining pattern that coincides with that of vinculin (a major cytoskeletal element in these cells). The CSAT antigen seemed to surround the membrane

focal contacts of vinculin, while also adjoining the loci of fibronectin (a major component of the extracellular matrix) contact.

Most recently, Dr. Horwitz has provided strong evidence that the CSAT antigen is a receptor for the two major extracellular matrix proteins, fibronectin and laminin. The three major lines of evidence are that the CSAT antibody inhibits adhesion, the antigen localizes in cell matrix adhesion plaques, and the purified antigen interacts with the putative matrix receptors. The antigen binds either of the matrix proteins with micromolar affinity. The CSAT antibody inhibits the formation of the antigen-laminin complex. Because of technical problems it proved more difficult to demonstrate a comparable inhibition of the antigen-fibronectin complex. However, a fibronectin-derived tetrapeptide, which binds to cells, does block complex formation.

The CSAT antigen has been purified and shown to contain three components, of about 160, 135, and 110 kD. The CSAT antibody will not bind to any of these bands. However, the SDS PAGE denaturation has not been reversed successfully, nor have the investigators been able to separate them under non-denaturing conditions. Dr. Horwitz and his colleagues have started collaborating with Dr. R. Hynes, of MIT, to clone the gene(s). A cDNA library from Avian fibroblasts has been screened using polyclonal antibodies against the CSAT antigen. So far, two clones have been isolated which synthesize 140 kD antigens that precipitate the antisera.

What effect does blockage of the CSAT antigen matrix attachment site have on the cells? The antibody was originally isolated because chick embryo myoblasts were seen to round and detach from substrate. The antibody can be shown to attach to CSAT protein in a wide variety of chick cells: chondrocytes; cardiac myocytes; skeletal myoblasts and myotubes; cardiac, skeletal, dermal, and tendon fibroblasts; and neurons. Whereas most cells detach from substrate after CSAT antibody exposure, cardiac fibroblasts seem surprisingly refractile, although the antigen is undoubtedly present. The antigen has been localized to all areas of the neuron. In the presence of antibody, axonal extension is blocked. Once the axon is extended, fasiculation is hastened by the antibody. The CSAT antigen appears to play a major role in axon routing and targeting. Such events have been much studied without a great deal of progress in obtaining a clear handle on any molecular event underlying these important growth processes. The identification of an antigen which is essential to normal nerve growth promises to open up a major new line of research.

"Cell Surface Lectins and Intercellular Adhesion"
RO1 GM 23547-09 (Rosen, S.), University of California, San Francisco

Cell-cell recognition is an important event in development, as well as in the normal physiology of the adult. Dr. Rosen at the University of California, San Francisco, initially approached the study of cell-cell interactions by investigating the role of cell surface lectins in intercellular adhesion, using the simple multicellular slime mold as an experimental model. He and others proposed that similar glycoprotein lectins play a role in mammalian cell recognition. In his most recent studies, Dr. Rosen has been testing that hypothesis with another interesting model for cell-cell recognition, that of lymphocyte recirculation.

During the process of lymphocyte recirculation, lymphocytes pass via the blood into secondary lymphoid organs such as lymph nodes and Peyer's Patches (gut-associated lymphoid tissues), where they encounter a specialized vasculature known as the postcapillary venules. These blood vessels are characterized by a cuboidal endothelial lining, and usually are referred to as high endothelial venules or HEV. A high proportion of lymphocytes carried by the blood into these vessels are able to bind to the apical surface of the specialized endothelial cells; the bound lymphocytes then migrate across the endothelium and thereby gain entry into the lymphoid organ. Here the lymphocytes encounter sequestered antigens. If the lymphocyte meets a cognate antigen, then an immune response can be mounted. If the appropriate antigen is not encountered, the lymphocyte will leave the lymphoid organ with the lymph, and eventually be carried back to the blood, whereupon it can reenter the same or another lymphoid organ. This process of lymphocyte recirculation serves to distribute immunocompetent lymphocytes throughout the body and thereby allows the body's array of lymphocytes to be brought into contact with the antigens that are sequestered in the various lymphoid organs.

Dr. Rosen has been investigating the molecular basis of this highly specific cell-interaction, focusing on two related aspects: 1) the identification of the lymphocyte surface glycoprotein lectin which binds to the HEV, and 2) characterization of the recognition unit for this lectin on the surface of HEV cells.

1) The Lymphocyte Surface Lectin--The lymphocyte lectin responsible for HEV binding has been isolated and appears to be an 80 kD molecule which selectively binds certain sugars, such as D-Mannose-6-phosphate (M6P). Dr. Rosen and his collaborators now have devised an artificial cell recognition system whereby fluorescent beads can be conjugated to a synthetic polysaccharide rich in M6P. These beads decorate lymphocytes, and their binding can be quantitated with a fluorescence activated cell sorter (FACS); the binding of the beads to lymphocytes demonstrates properties similar to those of lymphocyte binding to frozen sections of HEV (e.g., calcium dependence and trypsin sensitivity).
2) HEV Recognition Unit - Using frozen sections of lymph node HEV, Dr. Rosen has shown that treatment of the HEV with sialidase, an enzyme which cleaves sialic acid, prevents lymphocyte binding to the HEV. Thus, the interaction of lymphocytes with lymph node HEV may involve the recognition of at least two sugar moieties, D-mannose and sialic acid. These sugars may be present as distinct side chains of the receptor glycoprotein molecule.

A second interesting observation is that HEV from the intestinal lymphoid tissues, the Peyer's patches, still bind lymphocytes after sialidase treatment. These experiments suggest that sialic acid may be an organ specific recognition determinant involved in lymphocyte--HEV adherence for lymph nodes but not Peyer's Patches, whereas the 6MP sugar may be involved in binding to both kinds of HEV. Further studies are aimed at characterizing an acidic glycolipid, isolated from lymph node HEV cells, which may act as the attachment site for lymphocytes. Such acidic glycolipids also can be used to coat fluorescent beads which then bind avidly to lymphocytes.

Thus, these studies are defining both the nature of the glycoprotein molecules on lymphocytes, as well as the glycolipid binding sites on endothelial cells, and their distinct sugar residues which mediate a highly specific cell-cell recognition event of great physiologic importance.

"Control of the Cell Cycle"
R01 GM 24571-07 (Pardee, A.), Dana Farber Cancer Institute

As a cell grows and divides, it progresses through defined stages of a cell cycle; G_0 represents the resting stage, G_1 the growth phase, and S the phase in which DNA is replicated, prior to G_2 and mitosis. In normal cells, progression through the cell cycle is closely regulated, with many critical controlling events occuring in G_1. A cell in the G_1 phase of the cycle may either revert to a resting cell or progress to the later part of G_1, and then irreversibly to S, depending on a variety of factors.

The CMBD program supports several laboratories which are investigating various aspects of this central problem in cell biology. These laboratories are using various model systems, including prokaryotes, lower eukaryotes, amphibian cells and mammalian cell lines, to answer basic questions of growth control.

For example, Dr. Harvey Herschman of UCLA (GM 24797) is investigating the role of growth factors, such as epidermal growth factor (EGF) in events in G_1. Dr. Charles Stiles of the Dana Farber Cancer Institute (GM 31489), is looking at the relation of growth factors, such as platelet derived growth factor (PDGF), to proto-oncogene expression and function in growing cells. Several other laboratories, such as those of Dr. John Gerhart, Berkeley CA, (GM 19363), Dr. James Maller, U. CO, (GM 26743), and Dr. Thomas Laffler, Northwestern U., IL, (GM 29460), are working on factors which control the transition from the S phase to G_2 and mitosis, and regulation of key events in the mitotic cycle. CMBD supports a program project (GM 31286, Dr. Marc Kirschner at the U. California, San Francisco), which addresses multiple questions on the molecular biology of cellular growth control, in systems as diverse as the bacteriophage, developing drosophila and xenopus embryos.

Dr. Arthur Pardee and members of his laboratory, in a series of elegant experiments, have been investigating the molecular steps governing regulation of cell growth, progession through the cell cycle, and initiation of DNA synthesis in mammalian cells. Their approach is fourfold: to study 1) the role of growth factors in the G_0/G_1 and G_1/S transitions; 2) the regulation, expression, and function of proto-oncogenes, such as c-ras and c-myc in G_1; 3) biochemical events in late G_1 which are required for DNA synthesis; and 4) biochemical events associated with DNA synthesis at the beginning of S phase. The cell system that Dr. Pardee is using to study these questions is the 3T3 fibroblast cell line. A cloned subline of 3T3, called A31, has several features which make it particularly amenable to analysis of cell cycle regulation, i.e., 1) the cells can readily be synchronized, 2) addition of serum to the medium stimulates quiescent cells to proliferate, whereas serum deprivation causes actively growing cells to revert to the resting state, and 3) they can be transformed chemically to the neoplastic state (BPA31 cells).

In the past year, Dr. Pardee has made significant progress in dissecting the regulatory elements which control these aspects of cell growth in both normal and chemically transformed 3T3 cells. He has made major findings in the following three areas:

1) Growth factors—Three different growth factors are required for G_0/G_1 transition of A31 cells. In contrast, a single insulin-like growth factor (IGF) is required in late G_1 for initiation of DNA synthesis. The early growth

factors, contained in serum, cause transcription of new messenger RNA, whereas the late IGF acts post transcriptionally. Thus, when cells in G_1 are grown in a serum deficient medium, they regress to G_O, but upon addition of insulin, or IGF, they immediately proceed to the S phase. In contrast, continuously growing transformed BPA31 cells do not require IGF for DNA synthesis, indicating that they have lost this control point in late G_1.

2) Proto-Oncogene function--Expression of c-myc is greatly increased when quiescent cells are stimulated to grow. In contrast, in resting or growing transformed BPA31 cells, c-myc is constitutively expressed at the higher level, i.e., the BPA31 cells have lost cycle-dependent regulation of c-myc expression. However, expression of a second proto-oncogene, c-ras, is cell/cycle dependent in both growing A31 and BPA31 cells. Thus, the regulation of proto-oncogene expression and function is closely tied to the control mechanisms in the growth of normal cells as well as to aberrant growth in neoplastic cells.

3) The post-transcriptional progression from late G_1 to S phase appears to be dependent upon the presence of newly synthesized proteins. This progression can be prevented by growing cells in serum-deficient medium, whereupon they revert to a quiescent (G_O) state. The point at which they no longer will revert to a quiescent state, but are committed to DNA synthesis and progression through the rest of the cell cycle, is called the restriction point. It has been postulated that an unstable protein plays a key role in regulation of progression in normal cells, and that transformed cells have a protein variant that is more stable and leads to unrestricted growth. Dr. Pardee and co-workers have now isolated a 68 kD protein whose rate of synthesis and stability in transformed cells is greatly increased. This is the first such protein to be identified as a candidate for the restriction point control of cell growth.

In summary, Dr. Pardee is using a multipronged approach to determine the factors which regulate all phases of cell cycle progression, to study how these factors interrelate to one another, and to identify those aspects which are unique to normally regulated versus uncontrolled neoplastic cell growth.

Cellular and Molecular Basis of Disease Program

Program Organization and Professional Staff

Charles A. Miller, Ph.D.	Director
Bert I. Shapiro, Ph.D.	Deputy Director
Vivian F. Dickson	Special Assistant

Cellular Basis of Disease Section

Bert I. Shapiro, Ph.D.	Chief
Artrice V. Bader, Ph.D.	Program Administrator
Marion M. Zatz, Ph.D.	Program Administrator

Molecular Basis of Disease Section

Warren C. Jones, Ph.D.	Chief
Lee Van Lenten, M.D.	Program Administrator

Support Staff

Patricia P. Pluchino	Secretrary to the Director
Patricia Clark	Grants Clerk
Wilma Kline	Grants Clerk
Dana Quattlebaum	Grants Clerk
Judith Wilson	Grants Clerk

GENETICS PROGRAM

OBJECTIVES

The immediate goal of the Genetics Program continues to be the acquisition of knowledge about the fundamental processes and mechanisms of inheritance in health and disease. The ultimate objective of the research, research training, and resources supported by the Program is the prevention, therapy, or cure of human genetic disease. The Program staff considers the term "genetic disease" to encompass a broad spectrum of conditions. The term, as used in this context, includes not only those conditions with single-gene (Mendelian) modes of inheritance or chromosomal abnormalities, but also the more common degenerative conditions such as atherosclerosis and diabetes. Although the risk of occurrence of the latter kind of "multifactorial" conditions can be modified by environmental factors, a strong genetic component is generally involved. Despite recent triumphs in molecular genetics, methodology for the analysis of the genetics of common diseases continues to be an intractable and largely unsolved problem.

The staff of the Program seeks to promote constructive interactions with other NIGMS programs as well as with the other NIH Institutes. The research supported by the NIGMS Genetics Program tends to differ from research in genetics supported by many of the categorical (i.e., specific-disease-oriented) Institutes. For example, major emphasis is given to fundamental problems in genetics, such as the mechanisms of the transmission and expression of genetic information, and the control of such processes. Model organisms, such as microorganisms, fruit flies, corn, and mice are extensively employed in this kind of research. Thus, the research supported by the Program must be considered as a resource for generating both concepts and techniques that are important to, but transcend, concerns about any one genetic disease. For example, much of the research that led to recombinant DNA technology was supported by the Program. Thus, the work supported by the Program has had a large effect on the economic life of this country and, probably more significantly, also has provided scientists with tools to conduct types of research which were inconceivable only a decade ago. More recently, grantees of the Program developed new methodology for identifying the locations of deleterious genes on chromosomes, using restriction fragment length polymorphisms (RFLPs). This methodology has been responsible for a major breakthrough in research on Huntington's disease. It is now being intensively applied in research on cystic fibrosis, and it offers great hope in the diagnosis and ultimate eradication of other such serious diseases. Characteristically, the NIGMS Genetics Program has supported the research that has led to the concepts and methodologies, while subsequent research oriented toward specific diseases tends to be supported by one of the categorical Institutes.

ORGANIZATION AND STAFFING

The Genetics Program is organized into two sections that are designed to emphasize significant trends and perceived changes in the rapidly moving field of genetics: The Molecular and Medical Genetics Section and the Genetics of Growth and Differentiation Section.

Until summer 1984, the Molecular and Medical Genetics Section was headed by Dr. David Beck, who was also Deputy Director of the Genetics Program. Dr. Beck left NIH in July, 1984 to assume new duties at the Public Health Research Institute, New York, N.Y. Dr. Fred H. Bergmann, the Program's Director, is currently acting head of this section. In addition, the section includes Drs. Irene Eckstrand and James Cassatt (part time, in addition to duties with the Biophysics and Physiological Sciences Program). Dr. George Woolley, a long-time staff member in the Program, retired in January, 1985. Currently Ms. Dolores Lowery, a member of the MARC Program, is assisting us on a part-time basis particularly in the administration of National Research Service Award individual postdoctoral fellowships.

This section of the Genetics Program supports studies on the replication, recombination and repair of DNA, the mechanisms of mutagenesis, extrachromosomal inheritance, protein synthesis, human medical genetics, and cytogenetics and gene mapping. Until FY 1985, this section had also supported a number of studies on the physical and organic chemistry of nucleic acids. Research projects on these topics were transferred to the recently reorganized Biophysics and Physiological Sciences Program.

The other subdivision of the Genetics Program is the Genetics of Growth and Differentiation Section. The Chief of this section, who is also the Deputy Director of the Genetics Program, is Dr. Judith Greenberg. Dr. Jane Peterson (at NSF until June, 1985), Dr. Dorothea Miller (retirement effective November 1, 1985) and Dr. Barbara Williams assist in administering this component of the Program.

The Genetics of Growth and Differentiation Section supports studies on chromosome organization and mechanics, developmental genetics, mechanisms of transcription and the control of gene expression, rearrangement of genetic elements, and population genetics. During 1984, a new subsection of this program, on RNA processing, was created to reflect rapid developments in the field, as well as the relationship of RNA processing to post-transcriptional control of gene expression. In addition, another new subsection was designated to emphasize the emerging importance of studies on neurogenetics and the genetics of behavior in model systems.

In the past two years, the program has experienced a larger-than-normal turnover of support staff. Ms Dora Sullivan left in January, 1984, and Ms. Linda Virgens left in July, 1984. New support staff (since summer, 1983) are Ms. Martha Shanahan (joined in December, 1983), Ms. Jackie Whelchel (June, 1984), Ms. Lucy Clarke (left program in May, 1983, rejoined in March, 1985) and Ms. Maureen Smith (June, 1985). Ms. Mollie Hilty is now devoting a part of her efforts to the Biophysics and Physiological Sciences Program.

RESEARCH OVERVIEW

The diversity of the research supported by the NIGMS Genetics Program is best appreciated by a listing of the topics which we utilize to assign research projects to the two sections, and their respective subsections. Each of these subsections represents a portfolio of active grants, as well as applications currently under review, or under consideration for funding. Each of the professional staff members of the Program has responsibility for one or more of these programmatic sub-sections:

1. Molecular and Medical Genetics Section

21200 Replication, Recombination and Repair of DNA

Enzymes and mechanisms of DNA replication and repair; enzymology and genetic control of general homologous recombination; enzymatic modification of DNA; regulation of DNA replication, repair and homologous recombination; DNA replication associated with transformation and plasmid transfer; mechanisms of transformation, plasmid transfer and transfection.

21250 Mechanisms of Mutagenesis

Mechanisms of action of mutagens and carcinogens; genetics and biochemistry of mutation; interactions of mutagens with nucleic acids.

21280 Extrachromosomal Inheritance

Genetic aspects of mitochondria, plastids, kinetoplasts, and other extranuclear organelles; nuclear-cytoplasmic relationships in genetic expression.

21320 Mechanisms and Control of Translation and Assembly

Protein synthesis as a process, including initiation, elongation and termination; control of protein synthesis at the level of translation; structure, function, and isolation of cytoplasmic mRNA; effect of structure of mRNA on translation; interaction of mRNA with cytoplasmic proteins; structure and function of tRNA; structure, assembly and function of ribosomes and ribosomal components; structure and assembly of phages and viruses; function of phages and viruses when not serving as a model for other processes, e.g. transcription.

21400 Human Medical Genetics

Correlation of chromosomal anomalies with dysfunction; single gene defects including the genetic and molecular basis of metabolic errors; prenatal diagnosis and heterozygote detection; studies of polygenic inheritance in human genetic epidemiology; generalized approaches to therapy for genetic disorders; animal models of genetic disorders; pharmacogenetics where the emphasis is on genetics.

21480 Cytogenetics and Gene Mapping

Chromosome mapping; synteny and pedigree analysis for linkage; general cytogenetics; somatic cell genetics; molecular methods for gene mapping.

2. Genetics of Growth and Differentiation Section

21500 RNA Processing

All processing of RNA following completion of transcription of the particular gene region and prior to translational events in the cytoplasm: nuclear and analogous prokaryotic modifications of RNA, including methylation and capping;

cleavage of RNA; polyadenylation of RNA; formation, structure and function of hnRNA(P)s; splicing of RNA; formation, structure and function of snRNA(P)s; and transport of RNAs.

21600 Chromosome Organization and Mechanics

Arrangement of nucleotide sequences in chromosomes, and methods for detecting and analyzing unique and repetitive sequences; changes in sequence distribution and organization during development, differentiation and evolution; chromosome architecture and the interaction of structural components, including polynucleotides and histone and nonhistone chromosomal proteins; chromosomal mechanics, including segregation, crossing over, and chromosome rearrangements; general mechanisms of chromosome inactivation.

21700 Developmental Genetics

Genetic factors in germ cell formation and meiosis; genetic control of determination and differentiation of cell types; genetic control of the cell cycle and cell growth; mode of action of pleiotropic effectors that regulate cell differentiation and specialization; gene dosage and its expression during development.

21750 Mechanism of Transcription and Control of Gene Expression

Transcription as a process: structure and function of DNA-dependent RNA polymerases; repressor, operator, and terminator structure and function; RNA splicing when studied as part of transcription; action of pleiotropic effectors of transcription; differential gene transcription and its control except when the emphasis is on development; structure and organization of genes or complex loci as these relate to gene expression; and studies on gene rearrangement when emphasis is on the control of expression.

21780 Rearrangement of Genetic Elements

Studies of the transposition of genetic elements; insertion sequences and other specialized regions of recombination involved in gene transposition; mechanism and control of special types of gene rearrangement including site-specific and illegitimate recombination and gene splicing; integration and excision of viruses and other episomes.

21800 Dynamics of Genes in Population

Population genetics except those studies relating to medical genetics; genetic approaches to evolution including protein and nucleic acid polymorphisms in natural and laboratory populations; mechanisms of speciation; complex polygenic systems in populations; ecological genetics; environmental effects on genetic structure of populations; mathematical and biological model systems.

21850 Neurogenetics and the Genetics of Behavior

Genetic determinants of behavior patterns; genetic factors affecting learning and memory; action of genes in regulating differentiation of neural tissues and development of the nervous system.

RESEARCH TRAINING

In 1985, several important changes affecting the administration of training in genetics under the National Research Service Act took place. Following the retirement of Dr. George Woolley, who had administered postdoctoral fellowships in genetics since 1981, the Genetics Program moved toward greater staff involvement in training. Three professional staff members--Irene Eckstrand, Jane Peterson, and Dolores Lowery (part-time in Genetics Program)--cooperatively administer postdoctoral and senior fellowships (F32's and F33's). Judith Greenberg continues to administer the postdoctoral institutional training grant program.

The Genetics Program currently offers three training mechanisms--individual postdoctoral fellowships (F32), senior fellowships (F33), and postdoctoral institutional training grants. The major emphasis of the program is on individual postdoctoral fellowships, although a small number of senior fellowships to support retraining of experienced scientists are awarded each year. The postdoctoral institutional training program emphasizes research training for M.D.'s and Ph.D.'s in medical genetics. This program is highly selective and is likely to become more so as funds for training become more scarce.

Several trends in individual postdoctoral training are becoming apparent:

1. The following table provides the numbers of postdoctoral fellowships awarded by the Genetics Program, NIGMS, and NIH in 1980 and 1985:

	1980	1985
Genetics Program	340	237
NIGMS	676	407
NIH	1894	1774

In 1980, about 50 percent of the fellows supported by NIGMS were in the Genetics Program; 18 percent of all of the NIH individual postdoctoral fellows were supported by the Genetics Program. Since then, the number of fellowships has decreased sharply; further, the fraction of NIH fellows supported by NIGMS has dropped. Training in genetics, however, remains popular. As of August, 1985, the Genetics Program administered 237 individual postdoctoral fellowships--58 percent of the total NIGMS fellowship program and still 13 percent of the total number of individual postdoctoral fellowships awarded by NIH.

2. Approximately 4 percent of the individual postdoctoral fellows supported by the Genetics Program have M.D. degrees. Most of these M.D.'s also have Ph.D. degrees. By comparison, about 18 percent of all NIH fellows, who are supported by the categorical institutes, are M.D.'s.

3. The Genetics Program supports 13 institutional training grants which provide training for 51 postdoctoral trainees. Nearly half of the trainees have M.D. degrees.

4. Plant genetics is an increasingly popular area for postdoctoral training. About 15 percent of the individual fellows are engaged in research in which higher plants serve as model systems. It will be interesting to determine the impact of this training on future research trends.

The importance of postdoctoral training has increased with the need to master sophisticated genetic methodologies. The current objective of the training program is to train highly competent biomedical research scientists by providing excellent candidates with opportunities to develop their own research competence and creativity under the supervision of outstanding, experienced mentors. NIH-wide evaluations of postdoctoral training suggest this objective is being achieved.

CONTRACTS

A. Human Genetic Mutant Cell Repository (Contract NO1-GM-4-2100)

The Human Genetic Mutant Cell Repository, known widely throughout the genetics research community as the Cell Bank, is currently in its thirteenth year of operation. Over the course of these 13 years, the Cell Bank has continued to expand to meet the current needs of genetics researchers and to anticipate future needs. The Cell Bank is highly regarded for its service in providing well-characterized, contamination-free cell lines from patients with a wide range of genetic diseases and from unaffected family members. The collection now contains over 3500 cell lines. Many of these have characterized biochemical or cytogenetic defects, whereas others are from patients with genetic disorders in which the primary defect has not yet been identified.

Stimulated by the encouraging results obtained by correlating restriction fragment length polymorphisms (RFLPs) with the occurrence of a particular genetic defect in a large pedigree, geneticists are increasingly studying RFLPs as markers for a variety of disorders. In response to the need for large pedigrees for linkage analysis, the Cell Bank has acquired samples from multi-generational families with cystic fibrosis, diabetes, tuberous sclerosis, and ophthalmologic and connective tissue disorders. The Cell Bank also maintains extended pedigrees from normal individuals (the Utah collection) and from individuals with primary affective disorders (Old Order Amish collection) and Huntington's disease (Venezuelan collection and their normal family members). The Cell Bank has also recently initiated a small collection of somatic cell hybrids which will be of value for gene mapping.

In the past year, the Cell Bank shipped nearly 4000 cultures, to researchers in the United States and abroad. Among the most frequently requested cell lines are those from patients with Huntington's disease, xeroderma pigmentosum, cystic fibrosis, ataxia telangiectasia, and primary affective disorders. Cytogenetically abnormal cell lines, such as those containing trisomies, are also in great demand. In addition, cell lines from the extended Utah pedigree and from other apparently normal individuals are heavily utilized.

The Cell Bank publishes an annual catalog which serves as a valuable reference to investigators in genetics and cell biology. In each of the last several editions, the catalog has undergone improvements to make it increasingly use-

ful. For example, cell lines are cross-referenced by their McKusick numbers. This complements the listing of the Cell Bank's holdings which is included in the 1983 edition of Dr. Victor McKusick's Mendelian Inheritance in Man. Other additions to the catalog include reference pedigrees for the Utah, Old Order Amish, and Venezuelan pedigrees and diagrams of each chromosome showing translocation breakpoints and regions of monosomy, trisomy, or greater dosage corresponding to each chromosomally aberrant cell line in the collection.

Following a competitive review, the contract for the Human Genetic Mutant Cell Repository was again awarded in 1984 for five years to the Institute for Medical Research in Camden, NJ. The contract was recently modified to enable the Cell Bank to increase its emphasis on the characterization of cell lines at the molecular level.

The project officer for the Cell Bank is Dr. Judith Greenberg. She is advised by a group of highly dedicated scientists who perform an annual site visit to review the cell lines acquired during the preceding year, to recommend future acquisitions, and to establish general policies. Meetings to provide advice on specialized collections are also held periodically.

B. GENBANK®, THE GENETIC SEQUENCE DATABANK

In 1982, NIGMS established a computerized data bank of nucleotide sequences. Since that time, the data base has continued to grow dramatically and as of September 30, 1985, it included 5,204,420 base pairs comprising 5,700 sequences. The database is organized into 12 sequence groups to facilitate searching (primate, rodent, other mammalian, non-mammalian vertebrate, invertebrate, plant, organelle, bacterial, structural RNA, viral, bacteriophage, and synthetic/chimeric sequences). Each sequence is described by 20 data items covering size, base content, sites of biological interest, literature reference and source. The largest sequence to date in the data base is the complete sequence of the Epstein-Barr virus, composed of 172,282 base pairs.

GenBank® is available on-line via direct dial-up access or by Telenet and can be accessed from Prophet sites. Off-line distribution is primarily via computer readable magnetic tapes. Full releases of the data base are available quarterly with incremental updates available monthly. GenBank® is also distributed by sequence group on IBM-compatible floppy disks. A yearly hard-copy compendium is available as a four volume supplement to Nucleic Acids Research. The latter contains the combined contents of GenBank® and the European Molecular Biology Laboratory's Nucleotide Sequence Library.

GenBank® is supported under contract to Bolt Beranek and Newman Inc. (BBN) of Cambridge, MA. Data collection, entry and annotation is performed at Los Alamos National Laboratory under the direction of Dr. Walter Goad. Database management and distribution is handled by BBN under the technical direction of Dr. Wayne Rindone. Support for the data bank is provided by the NIGMS, the National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases, the National Institute of Allergy and Infectious Diseases, the National Cancer Institute, and the Division of Research Resources of NIH, the Office of Energy Research and the Office of Basic Energy Science of the Department of Energy, the National Science Foundation, and the Department of Defense.

The NIGMS is assisted in the scientific oversight of the data bank by curators appointed by NIGMS. These scientists are Dr. Dieter Soll of Yale University (tRNAs), Dr. Richard Roberts of Cold Spring Harbor Laboratory (adenovirus), Dr. Fred Blattner of the University of Wisconsin (lambdoid phages), Dr. Joachim Messing of the University of Minnesota (plants) and Dr. Elvin Kabat of Columbia University (immunoglobulins).

Enhancements to the data base this year have included distribution of the data base on floppy disks, and the provision of an on-line software clearing house to provide information on the availability of sequence analysis and manipulation programs. There are currently 25 entries in this auxiliary data base.

SMALL BUSINESS INNOVATION RESEARCH (SBIR):

As a result of the Small Business Innovation Research Act, agencies responsible for the funding of basic research are required to set aside a small percentage of their budgets to fund proposals submitted by small businesses. The awards fall into two phases. The purpose of phase I awards is to demonstrate the feasibility of the project. If phase I is successful, the applicant can apply for a phase II award, the purpose of which is to bring the project to the point of commercialization.

In August, 1985, the Genetics Program had three active phase II and two active phase I grants. Two of the phase II grantees were well along in their project and reported significant progress.

In one phase II grant Dr. Phillip Doggett of DNASTAR, in collaboration with Dr. Fred Blattner, President of DNASTAR, is developing software for the analysis of DNA sequences. The major advantage of the programs developed by DNASTAR is that they are designed to run on an IBM personal computer. The basic concept is to market an entire system complete with software. Included would be the IBM PC, a hard disk drive with sufficient capacity to store all the DNA sequences in GenBank® (both now and for the foreseeable future), a light box for reading sequencing gels, and the other customary peripherals (floppy disk drive, printer, etc.). Of special interest to NIGMS is that with this set-up an investigator would be able to manipulate the information contained in GenBank®, a facility funded largely by a contract from NIGMS. Included in the software developed are programs 1) for searching GenBank® for DNA sequences identical to or homologous with (to a specified degree) a given DNA sequence, 2) to determine the sequence of a fragment of DNA directly from a sequencing gel, and 3) to design an optimal DNA probe of specified length for a protein of known amino acid sequence. The project has made sufficient progress to allow test marketing of some of the software.

In another grant, Dr. Mickey Urdea of Chiron Corporation has refined existing technology to develop a new apparatus to synthesize short stretches of DNA oligonucleotides. Although of evolutionary rather than revolutionary design, the new synthesizer is a distinct improvement over models currently being marketed. Dr. Urdea reports that with his apparatus he can routinely synthesize oligonucleotides more than 100 bases long. Further improvement in this type of technology will be an important benefit to the many NIGMS grantees who are synthesizing oligonucleotides for structural studies of short stretches of DNA for use as probes to isolate selected genes, or as DNA segments for site directed mutagenesis.

In the third phase II grant, Dr. Robert Pergolizzi is developing biotin-labeled DNA prodes for the non-radioactive detection of DNA:DNA hybrids. This area is one of particular interest and importance as our ability to diagnose genetic disorders at the genomic level is growing rapidly. This type of development will permit the rapid transfer of the knowledge gleaned from basic science into the clinical realm.

RESEARCH HIGHLIGHTS, FY 1985

A. New Methodologies in Genetics

1. Overview: The Importance of New Methodologies to Molecular Genetics

In the period between 1950 and 1970, much of what we now call "mechanistic" or molecular genetics grew out of studies of relatively simple model organisms, such as the bacterium Escherichia coli and its viruses. In the seventies, and particularly since about 1980, these ideas have begun to be applied to research on higher organisms, such as the fruit fly, mouse, and man, organisms with far greater morphological, behavioral, and genetic complexity than E. coli.

The genome of E. coli consists of a sequence of nucleotides 4000 kb (kilobase pairs) long. By contrast, the human haploid genome is almost one thousand times larger (2.9 million kb). This correspondingly greater complexity strains the capabilities of current research methodologies to the limit. In fact, the rate of progress in molecular genetics is often very closely coupled to the development of new methodologies. The idea of using restriction fragment length polymorphisms (RFLPs) for human linkage assignments (Botstein, Davis, White, and Skolnick) has "rejuvenated classical genetic approaches and (has) already contributed to the chromosomal localization of loci associated with muscular dystrophy and Huntington's disease".* Nevertheless, the idea is only a few years old, and new methods which promise to extend its utility are of obvious importance (see section A2).

Although the use of recombinant DNA technology to obtain useful amounts of rare proteins for research and application is becoming commonplace, new methods are still being developed. The second highlight in this section details how an E. coli host can be coaxed to produce proteins which are extremely toxic and would normally be lethal to itself even if present in trace amounts.

Some thirty years ago, new techniques took perhaps a decade or more to become established as the standard tools of the average laboratory scientist. It is characteristic of the pace of modern genetics and its need for advanced methodology that the newest methods, such as RFLP analysis and pulsed field electrophoresis, have rapidly been adopted in scores of laboratories throughout the world. Characteristically, many of the developers of such methodologies are the grantees of the NIGMS Genetics Program--generally not because they are methodology-oriented but because the lack of suitable methodology is a bottleneck in their own research. It is thus most appropriate that a significant portion of the Genetics Program FY 1985 annual report is devoted to advances in methodology.

* P.N. Goodfellow, Nature 314:135 (85)

We have become so accustomed to thinking about the "revolution in genetics" as being due to recombinant DNA technology, now more than a decade old, that we may not be as aware of a second and more recent (three to four years old) "revolution" that has given us a radically new way of thinking about, and doing, genetics. This newer development is due to: (1) new technologies such as rapid DNA sequencing; isolation of large DNA fragments; the computer-assisted identification of control regions and "open reading frames"; many of which can be subsumed under the terms "chromosome walking" and "chromosome hopping"; (2) new concepts, such as that of domains in proteins as being informationally equivalent to exons in nucleic acids, differential splicing of mRNA precursors to get different gene products out of the same piece of DNA, and the conservation of important sequences throughout evolution--a new kind of comparative biochemistry; and (3) new resources, of which GenBank® and the Human Genetic Mutant Cell Repository, both administered by NIGMS, are good examples.

These newer concepts, techniques and resources have led to the development of a very new kind of genetics. One can now "map" a very complex trait about whose molecular biology we know essentially nothing--a complex developmental mutation in fruit flies, or an inherited neurological disease in humans--then directly analyze the regions of DNA at, or near, the genetic locus that seems to be involved. One then analyzes this region of DNA in a variety of ways and ultimately makes a sort of informed guess at how mutations in these sequences modify and disturb normal function. A vivid analogy to this kind of "reverse genetics" is to imagine a scientific milieu in which we know nothing about the molecular biology of hemoglobin, but can map a genetic mutation, which causes sickling of red blood cells and anemia. We could then isolate the piece of DNA responsible for these effects, and, with perseverance and luck, identify it as the gene for globin!

It is precisely in the area of correlating a genetic lesion with loss or disturbance of normal function that the two newer methodologies, described below, assume their importance. Anti-sense RNA can be used to "turn off" or inhibit a specific nucleotide sequence of DNA in vivo in order to determine how this might perturb normal cellular functions. In contrast, Dr. Leonard Lerman's method for producing many single nucleotide changes in a well-defined stretch of DNA offers the possibility of understanding, in great detail how each of hundreds of specific small changes in sequence alters function in a well-known protein such as globin, or a more poorly understood one such as a gene affecting development.

2. "Cloning of Certain DNA Sequences, Particularly Useful for RFLP Analysis, Requires Mutant E. Coli Hosts"
R01 GM 30467 (Wyman, A.), Massachusetts Institute of Technology

In 1980, a seminal paper by Botstein, White, Skolnick and Davis (Am. J. Human Gen., 32, 319) proposed that recombinant DNA technology might make possible the generation of a linkage map of the human genome. To develop such a map, one simply follows the co-segregation of restriction fragment length polymorphisms (RFLPs) as Mendelian traits in families.

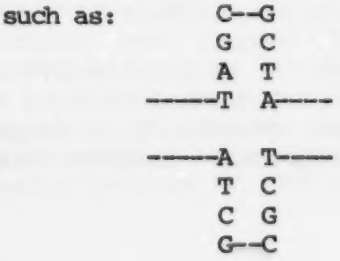
Usually, a sample of human DNA is digested with one of many possible bacterial DNA restriction enzymes. The huge numbers of DNA fragments that arise are separated by gel electrophoresis and transferred to a nitrocellulose filter where the DNA digest is hybridized to a radioactive probe. A pattern of bands, representing the limited number of DNA fragments that hybridize, is visualized by autoradiography.

Polymorphisms are merely genetic differences that can be attributed to variation in one genetic locus. Examples are the ABO blood groups, or certain enzyme variants. (The difference between a mutation and a polymorphism is that the former is a rare variant, the latter is a set of variants, each of which is relatively common in a population.) Because humans have two sets of genes, an individual can be either homozygous (one variant seen) or heterozygous (two variants seen). Prior to the development of the RFLP methodology, there were less than 100 polymorphisms in humans that could be easily detected and had enough allelic variation to be useful for genetic analysis. Probably the most useful of these continue to be the HLA polymorphisms, since there are many variants, at each of several genetic loci. Thus a random individual is much more likely to be heterozygous for the HLA markers, and thus, genetically informative.

Specific RFLPs are defined by the specific probe used to detect them, and also by the particular bacterial restriction enzyme used for the DNA digestion. The RFLPs that have been identified to date fall into two general classes. For some RFLPs, only two alleles (variants) are observed, resulting from either a single base pair mutation, or a single deletion or insertion event. At a different set of loci, multiple alleles are observed, as a result of multiple rearrangements. Like the HLA polymorphisms, this kind of locus is much more useful, since individuals are more likely to be heterozygous and genetically more informative. In 1980, Arlene Wyman and Raymond White discovered the first locus of this kind, which they called D14S1. Over 12 alleles were observed at this one locus, making it an outstanding marker for genetic analysis. It would be of tremendous utility if more such markers were found, since the more usual pair of alleles at RFLP loci is a very limiting factor in constructing and using RFLP maps. However, very few examples of this class of polymorphism could be defined by the existing probes.

Wyman recently attempted to clone several alleles of D14S1, and, in the process, discovered that the available probe they used hybridized to a region of human DNA outside of the hypervariable region. All attempts to clone the polymorphic region itself using a human genomic DNA library failed. She then recalled that, in 1984, the laboratory of Dr. Frank Stahl (GM 33677) reported that DNA containing inverted repeats of the the type -TAGCGCTA- which can loop out to -ATCGCGAT-

form structures such as:



cannot be cloned in bacteriophage lambda vectors unless recombination-deficient bacteriophage hosts are used.

Wyman, Wolfe and Botstein followed up on this clue. They packaged fragments of human DNA into lambda bacteriophage in vitro, and plated this mixture of viruses onto a strain of E. coli which had mutations in a number of loci needed for recombination (Strain DB 1170, mutant for rec B, rec C, and sbc B). They isolated 514 random phages containing fragments of human DNA, and tested them for their ability to grow on normal E. coli. Fully nine percent of these phages would not grow on wild-type strains of E. coli. Somehow, the presence of certain human sequences in the lambda phage is toxic to wildtype E. coli. (in the sense that they don't support phage growth) but not to the recombination-deficient strains. Further analysis of bacteriophages which could grow only on recombination-deficient mutants revealed that, as originally suspected, their complement of human DNA included inverted repeats 200-500 base pairs long. Furthermore, such sequences are often present many times in the loci specified by these probes. The presence of different multiples of such sequences could account for the large numbers of alleles seen in the loci specified by probes such as D14S1.

This finding may be very relevant to the fact that a number of specific sequences in mammalian DNA have proved difficult or impossible to clone. The most striking example is a region 3' to the alpha-globin locus. This region also shows a very high degree of polymorphism, similar to D14S1. Similar sequences are found in the 5' flanking region of the insulin gene. The results of Dr. Wyman's research suggest strongly that there is a connection between the inability to clone certain sequences in normal E. coli hosts, a high degree of polymorphism at the site, and the presence of different number of copies of an inverted repeat sequence. It seems also that a sizable fraction of the human genome contains such sequences. It may now be possible to clone many of them using recombination-deficient hosts. This class of sequences is highly polymorphic and thus is most useful for RFLP mapping. Thus these results are not only of intrinsic interest, but represent a significant contribution of bacterial genetics to the emerging technologies and strategies to map the human genome.

3. "Rare and Toxic Proteins Under Scrutiny"
R01 GM 21872 (Studier, W.F.), Brookhaven National Laboratory

DNA, the hereditary material, has two major functions--coding for all of the proteins needed for life processes and replicating itself in preparation for cell division and gamete formation. The replication process has proved difficult to study for two reasons. First, the analysis of replication-deficient cells is not easily accomplished because useful mutants are very rare; presumably, organisms with defects in the replication machinery seldom survive. A second approach to studying replication, biochemical analysis of replication proteins, requires that sizable quantities of the proteins can be isolated and purified. Replication proteins have not yielded rapidly to biochemical analysis because they are very rare--often fewer than twenty molecules are present in each cell.

The bacteriophage T7, whose genome of 39,936 bases codes for 50 proteins, has proven a useful system for solving these problems. T7 replicates by introducing a single molecule of DNA into a bacterial cell, whose resources are then diverted to the production of new phage. T7 is very efficient and prolific; as many as 250 new phage can be produced in 13 minutes.

T7 uses some proteins made by the E. coli cell to reproduce but also directs the production of specific T7 proteins using the E. coli cell's genetic machinery (ribosomes, transfer RNAs, amino acids). Inside the E. coli cell, T7's products perform a variety of functions. Some of the T7 gene products are involved in mobilizing the resources of the bacterial cell following infection. Others break down bacterial DNA and use it to make new copies of T7 DNA. Still others package new T7 DNA into viral particles. Many are specifically involved in replicating the viral genome.

Using recombinant DNA techniques, Dr. William Studier has been conducting systematic analyses of replication proteins and the genes which code for them. He initially devoted his energy to cloning and mapping the T7 genome, but as the cloned genes became available, Dr. Studier began detailed studies of specific genes, especially genes which code for replication proteins. All but two T7 genes have now been cloned and mapped and await further characterization.

Dr. Studier cloned the T7 genes by cutting each gene from the genome with appropiate restriction enzymes and inserting each into a suitable vector, such as the commonly-used plasmid, pBR322. The vector is capable of carrying the gene into a bacterial cell, where it can direct the production of a single protein. Two lines of experimentation are then possible:

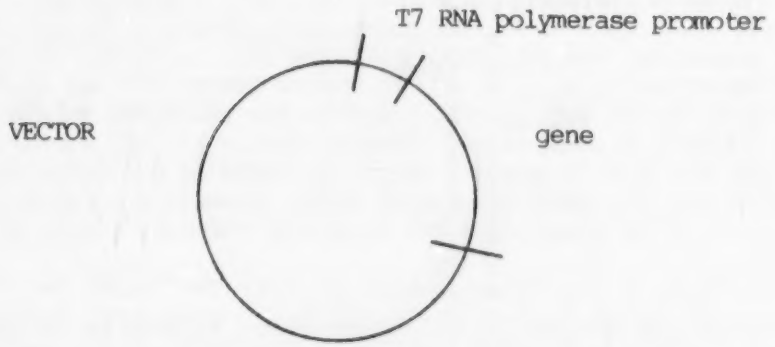
1. If the gene can be "turned on," the bacterial cell can become a factory for the gene product. This makes it possible to produce large amounts of a normally rare protein.

2. The existence of a bacterial cell that produces an essential viral gene product makes it possible to construct and maintain viral mutants that are totally defective in specific genes. The defective virus can grow on a bacterium that provides the missing gene product but not on any other.

Experiments using these strategies have been accomplished with great success in many laboratories; however, some proteins have proved impossible to produce. For example, some gene products are vital to cells in small quantities but may be toxic in larger quantities. Other are needed for T7 replication but are toxic to the E. coli cell. For example, T7 replication enzymes that nick DNA may degrade the E. coli chromosome before the bacterium can reproduce. Dr. Studier needed a way to turn on and turn off introduced T7 genes at will. Ideally, the gene under study could be turned off until a large number of copies in many cells were made. Then, all gene copies could be turned on simultaneously and large amounts of product would be rapidly produced.

In collaboration with Dr. John Dunn, Dr. Studier developed a simple and tractable system to study rare and toxic proteins of T7. Normal T7 phage produce their own RNA polymerase, which is capable of binding to a DNA sequence called the T7 RNA polymerase promoter. Only when the T7 RNA polymerase molecule binds to the T7 promoter will T7 genes be transcribed and expressed. The T7

promoter is highly specific for T7 RNA polymerase; RNA polymerase from other phage or bacteria is ineffective. Dr. Studier and Dr. Dunn managed to link the gene of interest, the vector to carry the gene into the cell, and the T7 RNA polymerase promoter, as follows:



The plasmid containing the structural gene and the T7 RNA polymerase promoter is inserted into an E. coli cell, where the plasmid rapidly begins to replicate. However, the structual gene is silent (not actively transcribed) until T7 RNA polymerase binds to the promoter and activates transcription. This system allows Dr. Studier to make many silent copies of a specific cloned gene, which he can then turn on at will by adding T7 RNA polymerase.

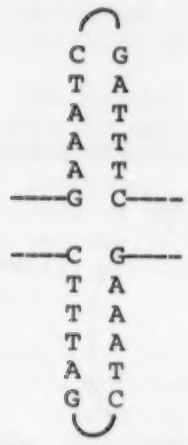
Dr. Studier is using this system to produce large quantities of T7 replication proteins, which he can then study biochemically. A good example of the success of this approach is his recent work on T7 gene 3. The product of gene 3 is needed for efficient T7 replication, for degradation of E. coli DNA during infection, and for genetic recombination. Large amounts kill host cells by degrading their DNA, so obtaining sufficient amounts to study has been difficult. Using the system described above, Dr. Studier successfully isolated enough gene 3 product for detailed study.

In collaboration with Drs. Bernard de Massey and Robert Weisberg at NIH, Dr. Studier found that gene 3 product is an endonuclease which cuts at specific structural sites in DNA. The scientists constructed a DNA substrate containing a palindromic sequence, that is, a sequence which is repeated in reverse such as the one shown below:

------GAAATC------GATTTC---

------CTTTAG------CTAAAG---

In vitro, such sequences readily form cruciform structures, in which the repeats pair as shown:



Cruciform structures are thought to be intermediate structures in the process of recombination. In the test tube, gene 3 product effectively removes cruciform structures from the DNA by clipping the DNA strand at each branch site. In the living cell, gene 3 protein is necessary for successful recombination and replication, presumably because it allows the stresses on the DNA to be relieved.

Although it is not the only system for isolating rare proteins, Dr. Studier's system may prove very useful to researchers in several areas. The T7 RNA polymerase promoter system has been used only for T7 and E. coli genes; however, it is likely to prove valuable for other groups of genes as well. The possibility of studying very rare proteins which affect important cell functions is very exciting. The goal of isolating and characterizing rare human proteins by this method is still in the future; however, studies such as Dr. Studier's suggest how it could be done. The next step toward that goal--expressing eukaryotic genes attached to the T7 RNA polymerase promoter--is now being undertaken. Dr. Studier's primary interest is still T7 DNA replication, but his work has led to development of a useful technique for scientists in many areas of study. It is a good example of how basic research in a seemingly narrow field can lead to important developments in many areas.

4. "Anti-sense Messenger RNA Makes Sense for Inhibiting Gene Expression"
   R01 GM 26176 (Weintraub, H.), Fred Hutchinson Cancer Research Center
   R01 GM 32921 (Melton, D.), Harvard University

Geneticists who want to know what product a gene makes or to understand how the activity of a gene is regulated frequently search for naturally occurring mutations or attempt to induce mutations in the gene of interest. This approach has worked remarkably well for Drosophila but has been less successful in organisms such as frogs or mice. Recently, investigators have begun to develop strategies that may circumvent this difficulty. Initial results from two laboratories suggest that it may soon be possible to turn off the expression of specific genes in cells of higher organisms, thereby mimicking the effect of a gene mutation.

DNA is a double-stranded molecule. From one strand (the coding, or sense, strand) is transcribed the messenger RNA (mRNA), which is then translated to produce a protein. The complementary strand, referred to as the anti-sense strand, is not ordinarily transcribed into mRNA because it does not have a promoter sequence, needed for proper transcription, in the correct location.

Drs. Jonathan Izant and Harold Weintraub reasoned that the anti-sense strand of mRNA should bind to the normal mRNA and block its ability to be translated into protein. This inhibition of translation should have an effect similar to that of a gene mutation which reduces or completely prevents synthesis of its protein product. To test this idea, it was necessary to construct a gene in which the anti-sense strand of DNA is able to be transcribed into mRNA. This was done by isolating the coding region of the cloned Herpes simplex virus thymidine kinase (TK) gene and reinserting it into a plasmid in the inverted orientation, with the promoter in front (at the 5' end) of the anti-sense DNA strand. When such a plasmid, which they named KF11, is introduced into a cell, the promoter directs transcription of the anti-sense strand.

Drs. Izant and Weintraub next injected copies of a plasmid containing the normal TK gene into mouse L cells which lacked a functional TK gene. The injected gene, if it is active, should direct the synthesis of the TK enzyme. Approximately 40 percent of the injected cells showed TK activity. However, when a large excess of KF11 plasmids containing the anti-sense TK gene was injected along with the plasmids containing the normal TK gene, there was a four-fold reduction in the number of cells expressing TK. Furthermore, cells still expressing TK had lower levels of activity. Control experiments ruled out the possibility that the promoter linked to the anti-sense gene competed with the promoter of the normal gene for factors required for transcription. Other studies showed that normal and anti-sense TK genes did not undergo significant recombination between themselves that might lead to inactivity of the normal TK gene.

Inhibition by anti-sense mRNA is of limited utility if it functions only for genes introduced at the same time as the inhibitor. It would be much more valuable to be able to inhibit the subsequent appearance of a gene product by preloading a cell with anti-sense mRNA. Drs. Izant and Weintraub therefore transfected KF11 plasmids into the cytoplasm of mouse L cells. Cells which integrated copies of the anti-sense genes into their chromosomes showed a reduced capacity to express normal TK genes which were injected into the nucleus later. This experiment showed that the anti-sense genes could become stably integrated into the host cells and presumably continue to make the inhibitory anti-sense mRNA.

Dr. Douglas Melton, using a slightly different approach, also demonstrated the ability of anti-sense mRNA to inhibit expression of a specific gene. His method involved synthesizing anti-sense mRNA *in vitro* and then injecting it into the cytoplasm of frog oocytes. A cloned piece of the *Xenopus* β-globin gene was linked to a promoter from *Salmonella* phage so that only the sense or anti-sense strand of DNA would be transcribed *in vitro*. Oocytes were injected with the resulting anti-sense mRNA for β-globin and incubated for five hours before receiving an injection of normal β-globin mRNA. When Dr. Melton measured the β-globin protein that was synthesized after a further five hour incubation, he observed that the anti-sense mRNA completely blocked translation. The

inhibition was specific, since general protein synthesis was not disrupted, and injection of anti-sense histone mRNA did not prevent synthesis of β-globin protein in cells injected with β-globin mRNA.

Dr. Melton further showed that the blocking of translation is due to hybridization, or binding, of the anti-sense β-globin mRNA to the normal -globin mRNA. The anti-sense mRNA appears to be relatively stable in the cytoplasm as long as the molecules are "capped", or contain a special guanine residue at the 5' end of the sequence.

Interestingly, for the anti-sense mRNA to block translation, it only needs to hybridize to the 5' end of the normal mRNA, the region where translation is initiated. Anti-sense mRNA which hybridizes to other portions of the mRNA is not effective in inhibiting translation. Taken together, these results indicate that anti-sense mRNAs block initiation of protein synthesis but not elongation of the protein.

The value of anti-sense mRNA as a tool for genetic analysis will become clearer when additional genes and other recipient cell types are examined. So far, however, the results from Drs. Izant and Weintraub and from Dr. Melton suggest that anti-sense RNA may provide new ways to analyze cellular functions associated with specific DNA sequences. For example, anti-sense mRNAs could help characterize genes that have regulatory roles during differentiation. Second, anti-sense RNA could help identify the functions of untranslated RNAs, such as small nuclear RNAs. Third, by blocking their translation, anti-sense mRNA could help explain the role of maternal mRNAs which are present in the cytoplasm of oocytes and which direct almost all protein synthesis during the earliest stages of embryogenesis. Fourth, anti-sense mRNA might be capable of inhibiting the activation of cellular genes induced by viruses, raising the possibility that specific anti-sense mRNAs could have therapeutic value in modulating the course of virally-induced infection or carcinogenesis. And finally, in conjunction with gene replacement therapy, anti-sense mRNA might be useful in shutting off expression of defective genes.

5. "A General Method for the Detection of Single Base Substitutions in DNA"
R01 GM 39095 (Lerman, L.S.), Genetics Institute; Boston, MA

Information about the variability of DNA sequences in eukaryotic genomes has been generated at a dramatic rate during the last few years. The advent of recombinant DNA technology has provided the tools to detect and analyse many genetic mutants and polymorphisms at the DNA level. A particularly informative technique has been the use of restriction enzymes to detect single base changes in the DNA sequence. This procedure relies upon the ability of a restriction enzyme to recognize and cleave DNA at a particular site. When a base change occurs within the restriction enzyme cleavage site, the enzyme is no longer able to recognize and cleave the DNA. The failure of the enzyme to cleave the DNA can be detected by gel electrophoresis. This technique has been used extensively to reveal numerous polymorphisms in the human genome. It is limited, however, because the base change must fall within the recognition sequence for a restriction enzyme and a radioactive probe must be available for the detection of each base change. Thus, for a disorder such as β-thalassemia, where 22 different single base changes have been identified, a large number of probes would be needed for diagnosis.

Another method to detect mutations or polymorphisms is to use S1 nuclease digestion of double-stranded DNA formed between a mutant and wild-type sequence. S1 nuclease is a single strand specific nuclease. Double-stranded duplexes formed between the wild type and mutant DNA will be single-stranded at the point of the base substitution. This is caused by the inability of the substituted base to pair with the wild type base. In this case, the mismatch is cleaved by the enzyme and the double stranded duplex separated by electrophoresis, revealing shorter DNA fragments in the mutant DNA. Application of this method is also limited because only certain base mismatches are cleaved by S1 nuclease.

Dr. Leonard Lerman at the Genetics Institute has been developing a method that may be more widely applicable for detecting single base changes than those mentioned above. This method is based on the knowledge that short duplexes of DNA formed from mutant and wild type DNA migrate more slowly during electrophoresis in a gel matrix containing urea and formamide than do native duplexes. The theoretical basis for the method stems from the differing stability of DNA base pairs in increasing concentrations of urea and formamide.

The dissociation of DNA strands at elevated temperature or in high urea concentrations is called "melting", a process during which the DNA strands begin to unravel. The point at which any given stretch of DNA melts is determined by the base composition of that stretch. Thus, different stretches of DNA will melt at different urea/formamide concentrations. When the duplex melts, it becomes more randomized in structure, and its progress through the gel in an electric field is slowed or stopped. Surprisingly, the substitution of a single base pair in the least stable part of the DNA fragment (usually the first 150 bp) introduces enough perturbation in the stability of the fragment to affect its electrophoretic mobility in a gel of increasing urea concentration.

In practice, DNA with a single base substitution is digested with a restriction enzyme, denatured and reannealed with a radioactively labelled piece of wild type DNA to form a heteroduplex. This DNA preparation is electrophoresed in a gradient of urea, where the mismatched piece of DNA begins to melt and is retarded in its migration through the gel. Every base substitution that Dr. Lerman has tested, up to base 144 or 146 in a 536 base pair fragment of DNA, has shown altered mobility under these conditions. The difficulty with detection of changes beyond base 144 appears to be due to the high melting temperature of the DNA in the more stable regions. Melting of the DNA in this region results in complete strand dissociation and therefore loss of resolving power of the gel.

These results suggest that a substitution is detectable if it falls in the less stable region of the DNA. Recently, Dr. Lerman, in collaboration with members of Dr. Tom Maniatis' laboratory, has used an elaborate tactic by which 95 percent of all DNA can be analyzed for single base substitutions. The single-stranded target DNA was subjected to conditions of severe chemical mutagenesis resulting in base substitutions in 10 to 20 percent of the treated DNA fragments. A primer was annealed to the DNA and a second DNA strand synthesized to form a DNA duplex. The double stranded DNA was cleaved by restriction enzymes, inserted into a DNA vector containing a "G-C clamp" and amplified. The G-C clamp is a guanosine-cytosine (G-C) fragment (which melts at a very high urea concentration) inserted next to the target sequence to prevent complete strand dissociation as the frag-

ment proceeds through the gel. After the plasmid was amplified, the target DNA plus G-C clamp was excised and subjected to electrophoresis on a preparative urea gradient gel to separate the mutant and wild type DNA fragments. The individual fragments were then isolated and sequenced. The application of this procedure to the β-globin promoter region resulted in the isolation and sequencing of a large number of fragments containing single base substitutions.

A unique feature of this work is the ability of urea gel procedure to recover many DNA fragments containing different single base substitutions rather than just a single fragment. Severe mutagenesis conditions producing a large number of molecules containing base changes are not generally used because of concern about the viability of the vector DNA in which target DNA resides. To avoid this problem, Drs. Lerman and Maniatis have excised the mutagenized target DNA after treatment and reinserted it into a viable vector for amplification. This strategy yields an array of fragments with single base substitutions at most sites along the DNA strand (called saturation mutagenesis).

Dr. Lerman points out that the separation technique works best when the mutagenized strand is reassociated with wild type DNA, producing a highly unstable mismatch. For example, a base substitution that has a minor effect on DNA duplex stability may be difficult to separate from the rest of the DNA. However, if it is duplexed with wild type DNA, the resulting mismatch at the base substitution will make the molecule less stable and it will be more clearly separated from the other DNA on the gel.

Dr. Lerman estimates that 95 percent of all DNA sequences should be accessible for analysis using this method or modifications of it. It will allow detection of mutations by a relatively straightforward procedure aiding in the diagnosis of various genetic diseases. In combination with the saturation mutagenesis treatment, scientists will be able to isolate large numbers of fragments containing different base substitutions for use in functional analysis of different DNA regions. Analysis of the whole human genome on two dimensional gels would be possible in theory if there were gels capable of resolving and displaying over a million spots! For the present, Dr. Lerman is concentrating his analysis on the globin gene.

B. Developmental Genetics

1. Overview

> Since differentiation is the basis of all multicellular life, its eventual elucidation is now a prime goal of modern biology. Until very recently, it was largely studied as an isolated subject, apart from modern genetic or biochemical ideas. Now, however, it is clear that the morphological tools of the classical embryologist cannot give satisfying answers. Instead, as in genetics, the fundamental answers must lie at the molecular level. The parallel with modern genetics may, in fact, be very close, since embryologists now believe that many of the basic control mechanisms that fix a cell's potential chemical reactions act at the level of the gene.

> Thus the recent advances which have made aspects of biochemistry and genetics indistinguishable may hopefully encompass the exciting aspects of embryology in the near future.
>
> J.D. Watson: Molecular Biology of the Gene, First Edition (1965)

Watson's prophetic words, written in 1965, and anticipating a new sort of embryology based on the molecular biology of the gene, began to come true in the eighties. In the examples below, two of the best model systems--nematodes and fruit flies--are utilized. Mutations are described which alter development both in space (i.e. homeotic mutants which determine the developmental fate of specific body segments in Drosophila) and in time (heterochronic mutants, which speed up or slow down the pre-programmed development of certain cell types during embryogenesis in nematodes).

The work described here depended heavily on the development of new scientific technologies, some of which were discussed in Section A. In particular, the nematode work rested heavily on new and sophisticated interference microscopy methods to trace normal or aberrant development of one egg through each of hundreds of different cell divisions, to the adult organism, composed of approximately 1,000 cells. In turn, the study of the developmental mutants of nematodes has provided hints about the mechanism of morphological change and morphological plasticity, not only in the development of the individual organism, but also at the level of evolution of diversity of biological form.

Drosophila developmental genetics initially relied on cytogenetics to roughly localize the position of various genes. The tools of chromosome walking and rapid nucleotide sequencing are now utilized to extend this work to a far finer level of resolution. An exciting spin-off, detailed below, is the emerging knowledge that parts of the developmental sequences studied in Drosophila have remained invariant during evolution in animals, and are seen in DNA from yeast, frogs, mice and humans.

Why are fruit flies and nematodes so successful as model systems for the study of behavioral and developmental genetics? It is because, from the geneticist's point of view, they are intermediate in size and genetic and morphological complexity between the "easy and less interesting" organisms, such as E. coli or yeast, and the "very difficult and very interesting" organisms such as mice and humans. They are complicated enough to exhibit a fairly complex repertoire of behavior and development. However, they are small enough to be grown in enormous numbers, so that interesting but rare mutants can be isolated. This intermediate level of complexity is reflected in the genome size of Drosophila. Each (haploid) cell of Drosophila contains 40 times as much DNA (165,000 kilobase-pairs, kb) as E. coli (4,000 kb), yet it contains 17 times less DNA than a human cell (2.9 million kb).

2. "Opening the Homeo Box in Vertebrates"
   R01 GM 09966 (Ruddle, F.), Yale University

The bodies of insects are divided into segments. Drosophila, for example, are composed of a head, three thoracic segments, and nine abdominal segments. It

has long been known that in Drosophila, as in other insects, mutations occur in which a part of a segment or an entire segment is transformed into a different segment. For example, certain mutants have their third thoracic segment transformed into a second thoracic segment. Since the second thoracic segment normally bears wings, a fly with an extra second thoracic segment will have two sets of wings. Mutations causing this type of anomaly are called homeotic mutations and define sets of genes (homeotic genes) which are required to specify the correct pattern of each segment during development of the fly.

From genetic crosses, it is known that the many homeotic genes in Drosophila are not randomly distributed on the chromosomes, but rather are arranged in two distinct clusters. One cluster, called the antennapedia complex, controls development of the anterior part of the animal, and the other cluster, the bithorax complex, controls the posterior part. A major advance in our understanding of homeotic genes occurred in 1983, when investigators in the laboratories of Thomas Kaufman (GM 24299) and David Hogness (GM 31409) mapped the two gene complexes. Following the mapping of these chromosomal regions, it was discovered that a highly conserved nucleotide sequence is present in as many as ten of the genes in the antennapedia and bithorax complexes. This sequence, named the homeo box, is 180 base pairs in length. From the DNA sequence of the homeo box, investigators deduced that it probably codes for a protein whose characteristics resemble those of DNA-binding proteins. The product of the homeo box may therefore interact with DNA to regulate the expression of other genes. This is in accord with the idea that homeotic genes serve as a set of switches which choose among a variety of cell fates. RNA transcripts and protein products of homeo box-containing genes have been found in cell nuclei in a stage-specific and cell type-specific manner in developing Drosophila embryos.

Once the Drosophila homeo box sequence was cloned, it was natural for investigators to search for similar regions in the genomes of other organisms. This is easily done by the method of Southern blotting, in which DNA from an organism is partially digested and the resulting pieces are electrophoretically separated on agarose gels. A replica of the gel is transferred to a nitrocellulose filter, which is then reacted with a radioactively labeled homeo box sequence. This sequence will only hybridize to pieces of DNA on the filter that contain similar sequences. Because the homeo box probe is radioactive, the pieces of DNA to which it binds can be detected. Using this approach, homeo boxes have been discovered in toads, mice, earthworms, chickens, and man.

Considerable attention has been focussed recently on determining the role of homeo boxes in vertebrates. The extreme conservation of homeo boxes among different species (70-90 percent homologous at the protein level) implies a common role in development. Many researchers in this area note that all the animals that appear to contain homeo boxes pass through a developmental stage when the body is composed of a linear series of segmental units. For example, all vertebrate embryos contain somites from which the basic organization of the skeleton, nervous system, and musculature is derived. Other researchers, however, suggest that the homeo boxes in vertebrates may code for a regulatory element that has nothing to do with pattern formation, or even that they may serve as a "tag" to target proteins to their correct location in a cell.

Among the investigators trying to determine the role of vertebrate homeo boxes are Dr. Frank Ruddle and his associates. They reasoned that if the location of homeo boxes relative to other genes on the chromosomes of different species is similar, their functions might also be similar, and this relationship might suggest a role for them. They therefore mapped the locations of homologous homeo box regions on human and mouse chromosomes, using as probes cloned human and mouse homeo box sequences along with long sequences that flank the homeo boxes. They mapped the mouse sequence to chromosome 11 and the human sequence to a region on chromosome 17 between bands q11 and q22. Other homeo boxes are present on other chromosomes, but these contain different flanking sequences and therefore are not detected by the particular cloned probes that were used.

Previous comparative mapping of mouse chromosome 11 and human chromosome 17 had revealed that several other genes are shared between those two locations. For example, the genes for galactokinase and thymidine kinase are closely linked to the homeo box loci in both mouse and human. These observations were independently confirmed by a second group of investigators.

Of interest in these findings is the fact that in the mouse the gene for short-tail is located on chromosome 11. Heterozygotes for the short-tail gene have skeletal abnormalities which are apparently the result of abnormal pattern formation. These include vertebral fusions, bilateral asymmetry of limb length, and an extra pair of ribs. Mapping at higher resolution will be required to determine if this gene defect, which resembles a homeotic mutation, is coincident with the homeo box locus on chromosome 11. If so, it would be tempting to speculate that genes associated with several human vertebral segmentation defects might be located near the homeo box on human chromosome 17.

3. "For Nematodes, It's a Matter of Time"
ROI GM 24663 and ROI GM 24943 (Horvitz, R.H.), Massachusetts Institute of Technology

Normal development of multicellular organisms from a single fertilized egg cell depends on the correct timing of specific developmental events, such as the division of cells, their differentiation to perform specialized functions, and their migration from one place to another. Development of a structure at an inappropriate time relative to the development of the rest of the organism could have serious consequences for the organism. Although one would expect developmental timing to be controlled by genes, there has been little direct evidence for the existence of such genes. Now, thanks to its special properties, the nematode Caenorhabditis elegans is providing information about genes that control temporal patterns of development.

C. elegans is a small, transparent roundworm with a life cycle of 3.5 days, and it consists, at maturity, of approximately 1000 somatic cells. As a result of painstaking work in several laboratories, every cell division, cell death, and cell migration that occurs to generate the embryonic, then the larval, and finally the adult form of C. elegans has been traced. These studies demonstrated that cell lineages are nearly invariant from one normal worm to another. Observations on living worms by Nomarski optics, therefore, allow comparison of development of individual cells in normal and mutant worms.

Several years ago, Dr. H. Robert Horvitz, in collaboration with Drs. John Sulston and Martin Chalfie at the MRC in England, identified a mutation in C. elegans which causes some somatic cells to repeat larval division patterns at a time when they should have stopped dividing to form adult-type cells. At that time, they named the mutant lin-4 (lin=lineage). It was recognized later that the lin-4 mutation is heterochronic, which means that it causes abnormal timing in the development of certain structures. Because one abnormality associated with the lin-4 mutation is defective egg-laying, Dr. Victor Ambros and Dr. Horvitz examined other egg-laying mutants to determine if they might also show alterations in developmental timing. Drs. Ambros and Horvitz have recently reported the discovery of other mutants, leading to the identification of three new heterochronic genes, lin-14 located on the x-chromosome, lin-28 on chromosome 1, and lin-29 on chromosome 2.

Mutations of lin-14 affect development in two opposing, and somewhat paradoxical, ways. Dominant mutations retard development by altering the fates of certain cells so that they repeat the lineage of their precursors instead of generating more advanced lineages. Recessive lin-14 mutations accelerate development in the affected lineages by causing cells to express the fates normally expressed by their descendants. Interestingly, there are two classes of recessive lin-14 mutations. The major class affects developmental events at all larval stages and in several different tissues, and a minor class affects a more limited group of lineages.

Drs. Ambros and Horvitz suggest that when the lin-14 gene product is synthesized at a high level, cells express early fates, and when it is produced at a lower level, they express later fates. This would imply that lin-14 activity decreases during development, causing cells to express first their early programs and then their later ones. Observations in other animals support this model for the action of lin-14. In the Mexican axolotl, a recessive mutation causes retarded expression of larval characteristics. The retardation can be overcome by exogenous application of thyroxine. Similarly, changing the levels of juvenile hormone at specific stages of insect development can cause heterochrony. Thus, thyroxine and juvenile hormone may mimic the effects of endogenously produced substances (heterochronic gene products) that regulate developmental timing.

Several alleles of lin-28 have been identified, which are recessive and cause acceleration of development. Like the minor class of lin-14 recessive mutants, these also affect only a limited number of lineages. Alleles of lin-29 which have been described are also recessive and retard development of selected lineages.

Drs. Ambros and Horvitz propose a model in which genes such as the major class of lin-14 recessive mutants are at the top of a developmental hierarchy and are involved in conveying temporal information to diverse cell types. Heterochronic genes, such as lin-29, which produce more specific effects, are lower in the hierarchy and under the control of the lin-14 genes.

The studies on C. elegans provide evidence that one gene, acting in some way to perturb timing of a variety of cell types, can cause an array of highly specific developmental defects. In this regard, heterochronic mutants can be compared to homeotic mutants, which also affect groups of specific cell lineages. Homeotic mutants, studied most thoroughly in Drosophila, cause one particular struc-

ture, such as a leg, to develop perfectly but in a location where a different structure, such as an antenna, should be. Whereas homeotic mutations affect spatial development, heterochronic mutations affect temporal development.

The discovery of heterochronic genes has implications for evolution as well as for development. Scientists who study evolution have proposed that new species may evolve because some aspect of development occurs too early or too late. The work of Drs. Ambros and Horvitz provides a genetic explanation for this form of evolution. For example, when one compares laboratory strains of C. elegans with related species of nematodes found in the wild, many differences are observed that resemble those caused by mutations of lin-14, lin-28, and lin-29. These include the absence of specific structures, the duplication of other structures, partial sexual transformation, and changes in the number of larval stages. This suggests that a mutation in a single heterochronic gene could, under appropriate conditions, cause evolution of a new species with many new characteristics.

Finally, it should be noted that C. elegans is proving to be a valuable model for studying development, genetics, and behavior. For certain types of studies it has advantages over Drosophila, the organism so frequently investigated as a model for higher animals. As the body of knowledge about C. elegans accumulates, this organism is likely to become increasingly useful in the laboratory. It's only a matter of time.

C. Replication, Recombination and Repair

1. Overview

The processes of replication, recombination, and DNA repair share many common properties and many of the same catalysts. They continue to emphasize to us that DNA is a molecule in continuous dynamic flux that reacts to the external environment, interacts with scores of highly specific DNA binding proteins, and changes its local conformation in response to all these events. There are complex devices in place to assure the fidelity of information carried by DNA. These permit proofreading of newly synthesized DNA and enzymatic correction of errors produced by environmental stresses. Similar machinery recognizes foreign DNA and degrades it. These are the devices that provide stability during times of flux, such as replication. Paradoxically, there are also devices such as genetic recombination that counter this stability by shuffling around DNA strands in order to assure genetic diversity. The number of proteins involved in these processes is very large, even in a system as simple as E. coli. The phenomena are often complex, requiring a good understanding of DNA physical chemistry and topology, enzymology, enzyme induction, and genetics. The identification of the many genes involved was often accomplished by very sophisticated genetic tools, prior to the actual isolation and purification of the host of proteins required. In what follows below, we have provided only a glimpse of this very active area.

Even an organism as simple as E. coli appears to have a number of options for DNA repair, depending on both the degree and type of damage from environmental insults as diverse as radiation and mutagenic chemicals. What is now obvious is that each repair pathway of an organism may require many enzymes-

topoisomerases, ligases, DNA polymerases, methylases, and endonucleases. Some of these enzymes are specifically induced as a result of the insult to the DNA, while others are already present for other functions, such as replication or genetic recombination. Furthermore, the local configuration of DNA--whether it is in the common B-form, the Z-form, or supercoiled--has an effect on the rate and extent of these reactions, and some of the enzymes listed above play a role in converting one form to the other.

Although much of the research on genetic recombination described here may seem somewhat esoteric, it has an extremely practical potential for the future use of DNA for therapeutic purposes. When foreign DNA is injected into a frog's egg, or taken up by a mammalian cell, for example, it usually integrates into the organism's chromosome in a random way. When the insertion is made in the middle of an essential gene, it can induce a deleterious mutation. A method which could assure integration at a very specific designated site would assure that DNA introduced in order to correct a genetic deficiency would not cause secondary damage of its own. The last item in this group (Section 7) indicates that efficient site-specific integration of foreign DNA is a foreseeable result of this research.

2. "The oriC Plasmid System Proves Its Usefulness"
R01 GM 07581 (Kornberg, A.), Stanford University

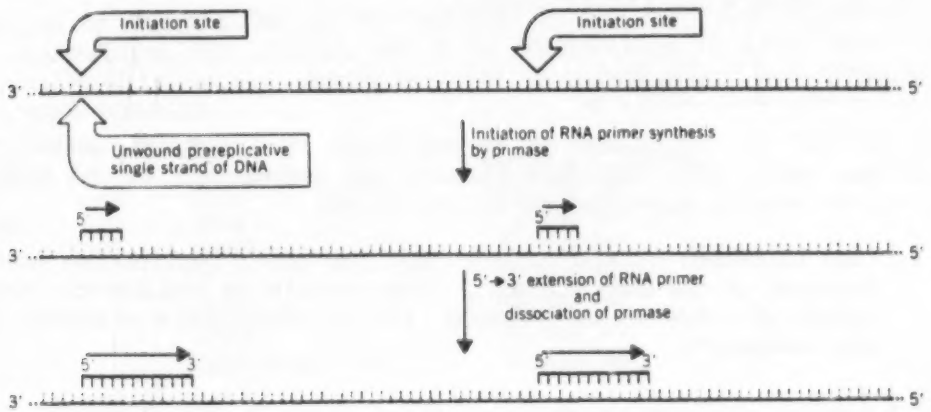
Replication of DNA is arguably a cell's most important function. Accurate and rapid replication is necessary to assure that the proper genetic information is passed from cell to cell and from generation to generation. Studies of DNA replication are difficult and slow largely because of the complexity of the biochemical reactions, the rarity of replication enzymes, and the inability to control the experimental system--the cell. Reproducing a complete replication system (that is, one which would both initiate and sustain DNA replication) in the test tube was, for many years, an important but elusive goal. In 1981, Dr. Arthur Kornberg announced the development of an *in vitro* replication system which promised to make studies of replication initiation and regulation more tractable. Recent studies in Dr. Kornberg's laboratory demonstrate that the optimism and enthusiasm which greeted the 1981 announcement were fully justified.

The system reported by Dr. Kornberg was based on a small, circular plasmid called pSY317. This plasmid was 12,000 nucleotide base pairs long and contained the origin of replication (*oriC*) from an *E. coli* chromosome. The presence of the 245 base pair *oriC* region made the plasmid functionally very similar to, but much smaller than, a normal *E. coli* chromosome, making initiation events relatively frequent and easy to study. Initiation of DNA replication in the original system required the addition of a crude cell extract and, surprisingly, ammonium sulfate, a chemical that is certainly not part of the *in vivo* replication system. Recent refinements of the system demonstrate that gyrase, an enzyme which affects the topological arrangement of the chromosome, can replace these components. The currently-used system achieves maximal DNA synthesis with a combination of highly purified enzymes, allowing for precise and controlled studies. Dr. Kornberg is now systematically studying the enzymology and genetics of two areas of DNA replication--prepriming of DNA synthesis and the role of DNA polymerase III holoenzyme. A brief overview of the steps in DNA replication will put Dr. Kornberg's work in its proper context.

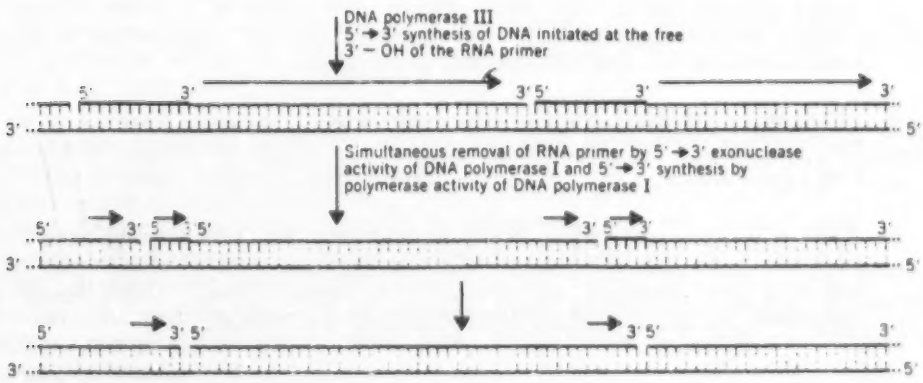
On the basis of earlier studies, scientists have separated DNA replication into four stages, as follows:

1. During prepriming, the oriC sequence is identified by specific proteins and the DNA strands begin to separate.

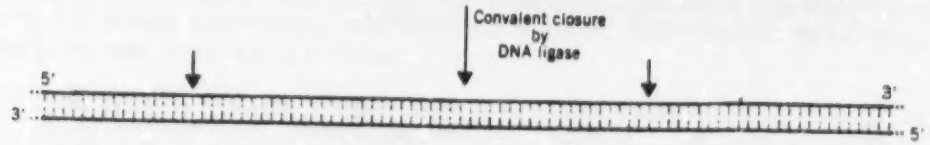
2. Once the DNA double helix has been opened at the origin of replication, priming begins. Enzymes called primases synthesize 10 60-nucleotide long RNA primers at several locations along the DNA strand. In the following diagram, the sites of primer synthesis are called initiation sites.


Initiation site
Initiation site
3'
5'
Unwound prereplicative single strand of DNA
Initiation of RNA primer synthesis by primase
5'→3' extension of RNA primer and dissociation of primase

3. Elongation involves the construction of a DNA copy of the DNA template. After dissociation of primase from the DNA strand, DNA polymerase III holoenzyme attaches to the DNA and begins DNA synthesis. DNA is synthesized in fragments between primer molecules. As the fragments are completed, DNA polymerase I simultaneously excises the RNA primers and synthesizes new DNA to fill the single-stranded gaps.


DNA polymerase III
5'→3' synthesis of DNA initiated at the free
3'—OH of the RNA primer
Simultaneous removal of RNA primer by 5'→3' exonuclease activity of DNA polymerase I and 5'→3' synthesis by polymerase activity of DNA polymerase I

4. During amplification, breaks between the fragments in the newly-synthesized DNA strand are repaired by DNA ligase.



Dr. Kornberg's recent work on prepriming reactions and the role of DNA polymerases III holoenzyme is straightforward and effective, making use of state-of-the-art biochemical and genetic approaches. Some of the recent results from Dr. Kornberg's laboratory follow:

Prepriming of DNA Replication

Replication of the oriC-plasmids is initiated by a complex series of steps. The process begins with the identification and preparation of the origin region and requires several enzymatic steps, as follow:

1. Identification of the origin site for other replication enzymes is a function of the dnaA protein. This protein is capable of binding to a number of sites on the plasmid, all of which share a common nine base pair sequence:

```
      C C
5'  TTAT CA A
      A A
```

The origin region is distinguished from the other binding sites by containing four repeats of this sequence. At oriC, about 250 base pairs are covered by complexes of 20 to 30 dnaA protein molecules; at the other sites, only 100 base pairs are covered. Identification of the origin region appears to be a function of the size and perhaps the conformation of the DNA/protein complex.

2. Initiation of replication at other dnaA binding sites is also suppressed by the action of accessory proteins, such as RNase H and protein HU. The mechanics of this suppression are not yet understood.

3. dnaB and dnaC proteins form a complex and are transported to the dnaA protein. It is possible, but still unproven, that the dnaC protein targets the complex to dnaA. It does appear that dnaC protein is released after the dnaA/dnaB complex is formed.

4. dnaB protein and gyrase open and unwind the DNA strand in preparation for binding by the replication proteins. DNA is normally helical and tightly supercoiled. Separation of the DNA strands at the origin could cause the remainder of the molecule to wind more tightly. The complementary actions of dnaB protein and gyrase prevent this from happening. dnaB protein separates the strands at the origin region while gyrase acts as a swivel to relieve the tension on the DNA molecule. (Work in other labs has shown that gyrase clips one DNA strand, allows the strands to pivot around each other, and then rejoins the strands.)

5. Once the origin is identified and unwound, the priming reactions begin.

DNA Polymerase III Holoenzyme

The replication of E. coli DNA is carried out primarily by DNA polymerase III holoenzyme (holoenzyme, for short), which is composed of at least seven subunits. Recent studies in Dr. Kornberg's lab have identified the genes which code for several subunits.

Subunit	Gene	Subunit Size
alpha	dnaE	140,000 MW
epsilon	dnaQ	7,000 MW
beta	dnaN	38,000 MW
theta	?	10,000 MW
gamma	dnaC	50,000 MW
delta	dnaX	28,000 MW
tau	dnaX+dnaC (?)	78,000 MW

Four of these genes have now been cloned, opening the way for producing large quantities of their rare products.

Dr. Kornberg has also been involved in studying the dynamics of holoenzyme as it moves along the template. Results so far indicate the following:

1. Holoenzyme begins synthesizing new DNA at the 3' end of the primer, adding nucleotides and moving along the DNA template. The DNA/enzyme complex is very stable; holoenzyme can remain bound to DNA for well over 15 minutes.

2. Holoenzyme is also efficient; it continues to add nucleotides to the growing strand until it encounters the next primer. Virtually no bases of the template are left unpaired.

3. Until recently, the behavior of holoenzyme upon encountering 5' end of the next primer was unknown. Studies of other DNA binding proteins suggest several possibilities. The enzyme might dissociate from the DNA and "search" randomly or in a directed way for the 3' end of the next available primer. Alternatively, holoenzyme might bind to the DNA/RNA duplex without dissociating from the template. Dr. Kornberg and his colleagues have now established that holoenzyme dissociates briefly (under 1 second), and then reattaches to the 5' end of the DNA/RNA primer.

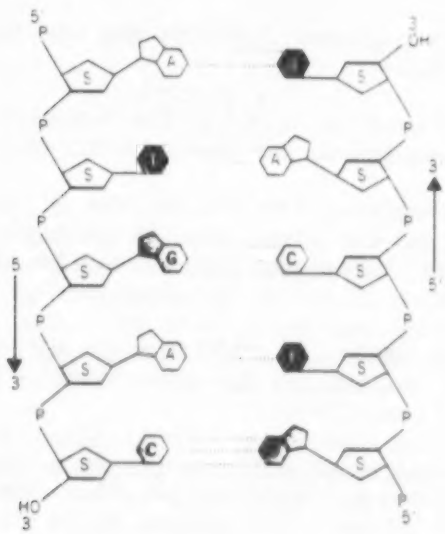
4. Holoenzyme diffuses along the duplex DNA/RNA region until it encounters the 3' end of the primer. It is then ready to reinitiate DNA synthesis.

ATP plays a critical role in this process; it is responsible for the stability of the DNA/holoenzyme complex, apparently enabling holoenzyme to grip the template DNA strand. Dissociation of holoenzyme from the template requires breaking this firm bond and is also ATP-dependent; however, movement of holoenzyme along double-stranded regions does not require ATP.

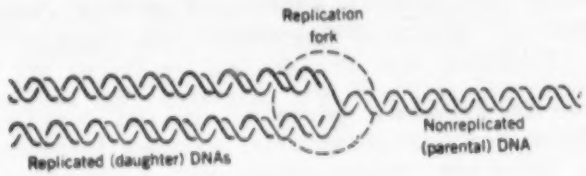
Models of enzyme action at the replication fork will undoubtedly develop as more data are accumulated. In this regard, the usefulness of the oriC plasmid can hardly be overestimated. Its development has opened the way for laboratories around the world to conduct controlled, detailed experiments which are revealing the enzymology and genetics of DNA replication.

3. "Artificial Gene Aids Crystallographic Studies of a Replication Enzyme"
ROI GM 28550 (Grindley, N.), Yale University

DNA is a molecule with a mission - making copies of itself. Elaborate biological processes and structures make this possible, assuring that DNA is faithfully duplicated from one generation to the next. DNA is a double-stranded molecule, composed of two long backbones made of alternating deoxyribose sugar and phosphate units and held together by nucleotide bases. The strands are said to be antiparallel because they are in opposite orientation, as shown below:

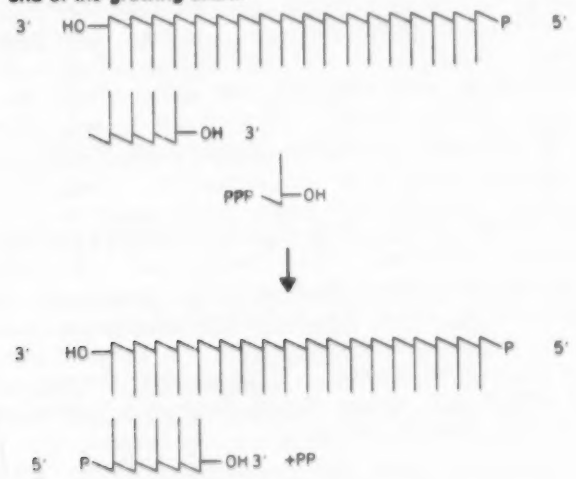


The sequence of the nucleotide bases--adenine (A), thymine (T), guanine (G), and cytosine (C)--comprises the code for enzymes, structural and transport proteins, ribosomal and transfer RNAs, and other cell components. During replication, the two strands separate and complementary strands are synthesized. The two new DNA strands are thus composed of one parental strand and one newly-synthesized strand.



The replication of DNA requires many enzymes, such as topoisomerases, single-strand binding proteins, and polymerases. DNA polymerase, one of the most complex of the replication enzymes, performs the most basic function--that of transferring the genetic code from the parent molecule to a new, complementary strand. Although the first DNA polymerase, pol I, was discovered in the mid-1950's by Kornberg (GM 07581), the enzyme remains poorly understood. Scientists know that DNA polymerases, including pol I, add nucleoside 5'-triphosphates to the 3' end of a growing DNA strand. The enzyme must also be able to "read" the template strand so that the new strand is complementary.

DNA synthesis occurs by addition of nucleotides to the 3'-OH end of the growing chain.



DNA polymerase I (pol I) is not an extraordinarily large enzyme (it contains about 1000 amino acids), but it replicates DNA with remarkable efficiency and speed. In normal conditions, it constructs complementary DNA strands at the rate of 667 nucleotide bases each minute with an error rate of about 1 in 10^6 to 8 base pairs! The physiological roles of pol I in living cells are poorly understood; most of our understanding of the enzyme comes from in vitro studies. These studies clearly show that pol I has three distinct activities--synthesis of new DNA strands (polymerase activity), repair of errors made during DNA synthesis (5' to 3' nuclease activity), and excision of large distortions of the DNA molecule (3' to 5' activity).

These three functions reside in different parts of the pol I molecule, making it a useful model for studying how multifunctional enzymes interact with their substrates. Unfortunately, progress toward obtaining a detailed structural description of pol I has been slow, in spite of the availability of clones which can be used to produce large quantities of the complete enzyme.

The difficulty lies in the size of the molecule (MW=109,000). No molecule of pol I's size and complexity has yet been solved crystallographically. Fortunately, the pol I molecule can be split into two enzymatically active fragments. The small fragment (MW=36,000) contains only the 3' to 5' nuclease activity; the large fragment (MW=76,000), called the Klenow fragment, contains the polymerase and editing functions. Because the Klenow fragment is only two-thirds as large as the intact enzyme but retains the most commonly studied activities, it has become an extraordinarily useful research tool. Furthermore, its size permits high resolution structural studies.

Dr. Nigel Grindley has recently used recombinant DNA techniques to construct an artificial gene which can produce large quantities of the Klenow fragment. As it turns out, the first third of the normal pol I gene codes for the small fragment of the enzyme, while the remaining two-thirds codes for the Klenow fragment. Dr. Grindley cloned the complete pol I gene, then positioned strong promoter sequences and start signals at the beginning of the Klenow fragment coding sequence. The non-Klenow portions of the gene were deleted, and the modified gene was inserted into a vector. Placed in a suitable environment, such as E. coli, the artificial gene produces large quantities of normal Klenow fragment.

The Klenow fragment was provided to Dr. Tom Steitz and his colleagues at Yale, who began the crystallographic studies. They determined that the Klenow fragment itself is composed of two domains. The smaller domain is about 200 amino acids long and contains a binding site for a deoxynucleoside monophosphate and a divalent metal ion. This region probably makes up the active site for the editing function of the polymerase enzyme. The larger domain which performs the polymerase function is about 400 amino acids long and contains a deep cleft of the right size to hold a DNA double helix.

Grindley and Steitz went on to study the Klenow fragments of DNA polymerase mutants. Two mutants known to have defective interactions with the DNA molecule have amino acid changes on the surface of the deep cleft. Data now indicate that actual synthesis of new DNA segments takes place along one wall of the cleft in the large domain. In the normal molecule, this position is a long distance from the active site for the editing function, raising the possibility

that the DNA polymerase molecule may undergo radical position changes as it moves from the polymerase to the editing mode. The crystallographic studies also indicate a structure which may form a lid to the deep cleft after the DNA is bound. This structure probably inhibits dissociation of the DNA/enzyme complex.

DNA polymerase is one of the most interesting and important molecules in cells. It is responsible for replicating and repairing DNA; it is also an excellent model of the enzymes which interact with DNA and which have many functions. These studies by Grindley and Steitz are providing a vivid picture of how DNA polymerase and DNA interact and how DNA polymerase carries out its many functions.

4. "RecA Protein and the Mechanism of Recombination"
ROI GM 11014 (Howard-Flanders, P.), Yale University

In 1911, Thomas Hunt Morgan published a classic paper in which he hypothesized that genes are arranged in a linear sequence on chromosomes and that crossing over (the exchange of genetic material) occurs between homologous chromosomes during meiosis. Morgan thought that crossing over was caused by mechanical forces acting on the chromosomes as they twisted around each other during synapsis and that it was, therefore, an aberrant event. The importance of crossing over was not lost on Morgan, however, and with his student, A.H. Sturtevant, he developed the technique of mapping genes onto chromosomes by using crossover frequencies to estimate relative distances between genes.

Today, we know that crossing over is only one subset of a larger group of phenomena called recombination and that recombination occurs, not by mechanical breakage, but by enzymological processes. Many scientists now believe that recombination, far from being an aberrant event, is necessary for proper assortment of chromosomes into daughter cells during cell division. In fact, recombination processes are important at all levels of genetic organization. At the molecular level, recombination is a tool to facilitate fine structure mapping of mutations within genes. Transposition of mobile genetic elements, insertion of viral genomes into host chromosomes, and repair of genetic damage require recombination processes. In addition, the great diversity of life on earth is due to biological processes that generate genetic variation. Recombination of genes into new configurations is an important element of evolution.

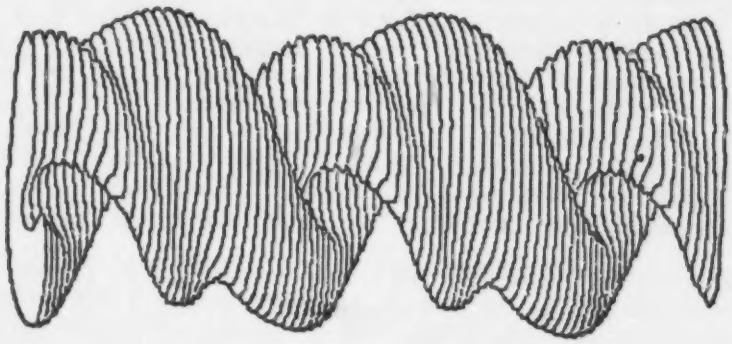
Understanding the mechanism of recombination has been difficult, largely because the organisms most suitable for genetic analysis, such as fungi, are not the best for molecular analysis. Conversely, E. coli is an excellent model for many molecular studies, but because recombination here occurs between phage and E. coli chromosomes and because phage life cycles are diverse and often bizarre, it is impossible to generalize to eukaryotes. Despite this handicap, research on the biochemistry of recombination has been proceeding.

In 1965, Clark and Margulies isolated E. coli mutants that were unable to undergo recombination during conjugation. The defective gene was identified and called recA. Paul Howard-Flanders has been closely involved in studying the recA gene product since its discovery and has recently proposed a model for the interaction of recA protein and DNA.

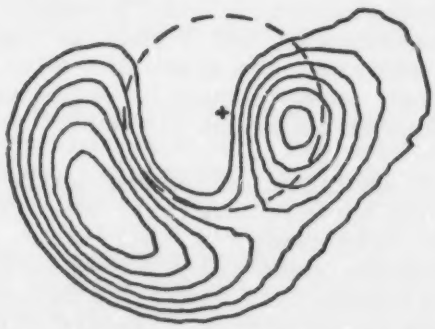
RecA Spiral Filaments

X-ray diffraction studies of recA protein crystals show that the molecules are arranged in long helical strands with six recA monomers in each turn of the helix. In solution, the protein binds to both single and double stranded DNA to form long spiral helices with 18.6 DNA base pairs and 6.2 recA monomers in each turn of the helix. The DNA in this complex is stretched by about 50 percent compared to normal, unbound DNA.

Dr. Howard-Flanders has recently completed research making use of imaging techniques to study recA/DNA complexes. Uranyl acetate is a commonly-used negative stain because it is excluded from protein structures. Thus, electron micrographs of stained material reveal the outline of a protein against a stained background. Dr. Howard-Flanders stained recA/DNA spiral filaments with uranyl acetate and used a Fourier transformation procedure to construct a 3-dimensional image of the recA spiral filament. The following picture, drawn by computer, shows that the helical filament contains a deep groove.



A computer-drawn cross section of the recA filament confirms the existence of a deep groove.

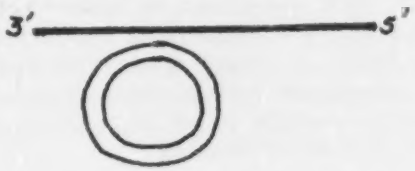


Dr. Howard-Flanders has calculated that the stretched DNA must lie within a 40°A diameter circle in order to avoid disrupting any bonds. If the DNA were to lie within the deep groove of the recA filament, it would meet this constraint.

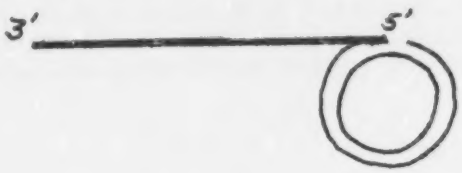
Pairing and Strand Exchange Within the RecA/DNA Complex

The most common biological models used to study recA are very simple systems, usually consisting of a single-stranded DNA (ssDNA) molecule and a homologous double-stranded (duplex) DNA molecule. If recombination occurs, the ssDNA will be substituted for one of the strands on the duplex DNA. Assaying for recombination is difficult, however. Another system that is easy to study consists of linear, ssDNA and a homologous circular, duplex DNA. Kevin McEntee (GM 29558) has used this system to determine the steps involved in recA-mediated recombination.

1. RecA protein binds to ssDNA. This reaction is facilitated by the addition of single-strand binding protein. (On the pictures below, the recA protein subunits are not illustrated).



2. Although recA protein alone binds slowly to duplex DNA, the recA/ssDNA complex readily pairs with duplex DNA. One strand of the duplex circle is nicked.



3. The ssDNA replaces one strand of the duplex DNA. This reaction requires ATP and proceeds in a 5' to 3' direction along the ssDNA.

4. The result is a single-stranded linear molecule of DNA and a nicked circular duplex.

In many forms of recombination, such as crossing over, both DNA molecules are double-stranded. This is more complex because four DNA strands, not just three, are involved. Dr. Howard-Flanders has found that binding of the recA protein to the DNA requires a gap (a single-stranded region) in one of the four DNA stands. After binding to the single-stranded region, the recA protein promotes strand exchange between the two duplex molecules.

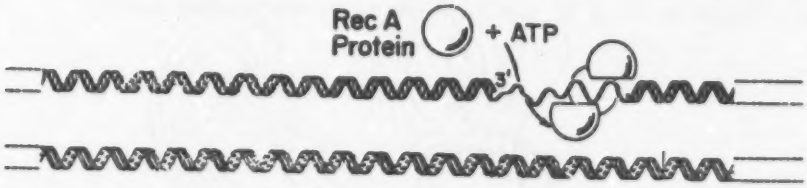
DNA Binding Sites in the RecA Protein

One the basis of the sizes of DNA and the recA protein, Dr. Howard-Flanders has proposed that there are two DNA binding sites in the deep groove of recA monomers. In addition, he has identified recA mutants which can bind ssDNA, but not duplex DNA. This supports the idea that each recA monomer has two DNA binding sites--one for ssDNA (or gapped double-stranded molecules) and one for duplex DNA.

Model for DNA/RecA Interaction

Based on these data, Dr. Howard-Flanders has proposed an interesting and provocative model for how recA protein facilitates recombination. The following series of illustrations shows the main features of the model:

A



RecA protein binds to a gap in one of the DNA molecules,

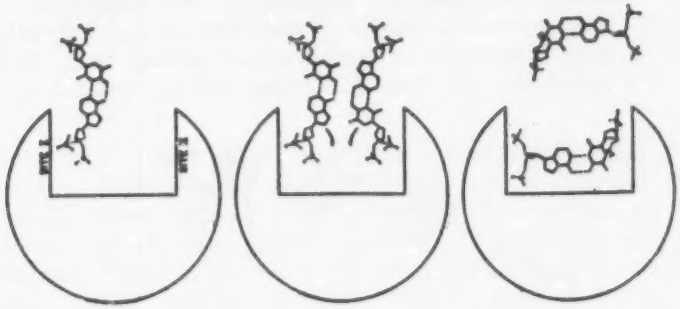
B

RecA promotes contact between the ssDNA and the duplex DNA molecules. As monomers of recA coalesce around the DNA to form a spiral filament, the ssDNA associates with the duplex DNA forming a 3-stranded DNA molecule.

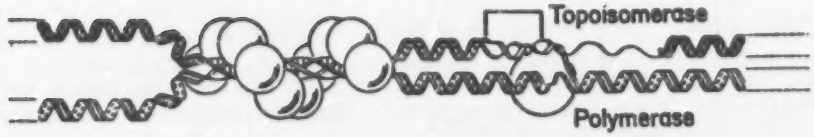
C

The filament winds around the 3-stranded DNA molecule.

At the end of the gap, recA will bind to the remaining DNA strand forming a 4-stranded DNA complex. In this configuration, the bonds that hold the DNA strands together are easily broken, and the nucleotide bases can rotate freely. Thus, strand exchange is easily accomplished, as shown below:



D



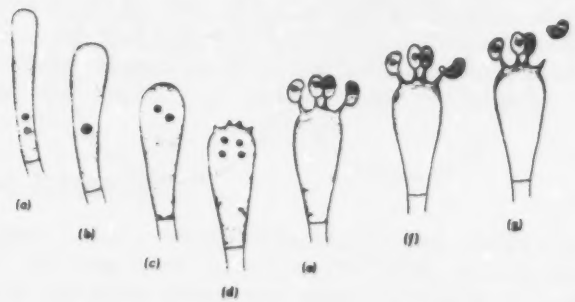
Topoisomerases relax the DNA helices into their normal configuration, and DNA polymerase repairs the single-strand gap.

This model is an important advance in understanding genetic recombination. Although many details remain unresolved, the basic working model is gaining acceptance as the first plausible molecular model of recombination. The intellectual distance between Morgan's first explanation of genetic traits and Dr. Howard-Flanders' model is great and highlights the remarkable advances in genetics over the past 73 years. It is satisfying to know that state-of-the-art technologies are finally answering the questions and explaining the data that T.H. Morgan and his colleagues produced.

5. "A New Twist in the Z-DNA Story"
R01 GM 27103 (Holloman, W.), University of Florida

Genetic recombination is a common process during which DNA strands exchange genetic material. Recombination is found in all organisms, from bacteria to humans, and is now believed to be necessary for proper cell division and cell differentiation. Until recently, scientists believed that recombination was caused by mechanical forces acting on the chromosomes as they twisted around each other. We now know that special enzymes carry out the process and that cells without these enzymes or with defective enzymes are abnormal in many cell functions. Recombination is important in the transposition of genetic elements, insertion of viral genomes into host chromosomes, repair of genetic damage, and the assortment of chromosomes into daughter cells during cell division. In addition, the great diversity of life on earth is due to biological processes that generate genetic variation. Recombination of genes into new configurations is an important element of evolution.

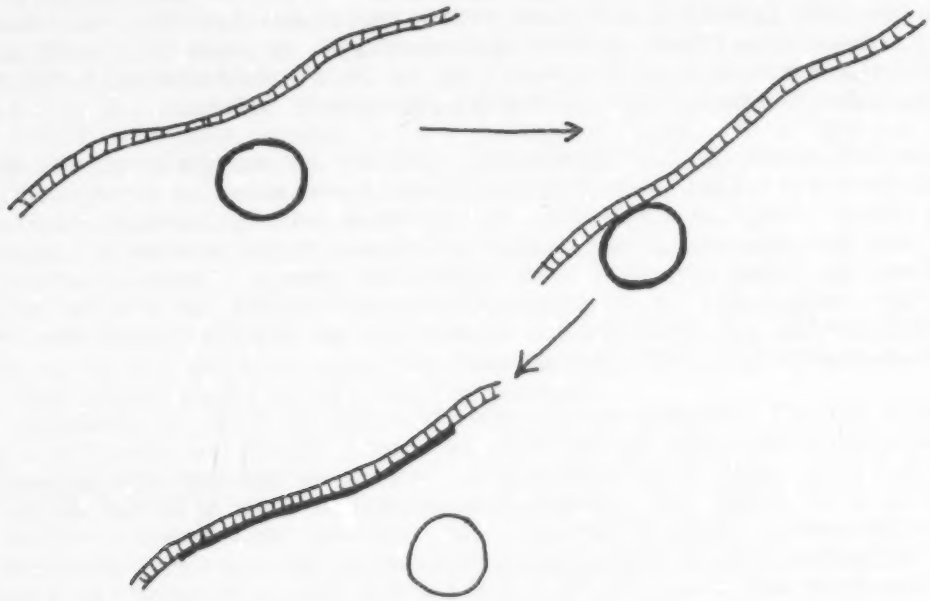
Dr. William Holloman has been studying genetic recombination in Ustilago, a yeast known better as corn smut. Because recombination occurs mainly during meiosis, the process is usually studied by examining the products of meiotic cell divisions. Ustilago has proved a useful model for studying recombination because all products of a given meiotic or gamete-producing cell division can be gently picked from the surface of the dividing cell, called a basidium.



Production of Gametes in Basidium of Ustilago

Recombination in Ustilago is carried out by the recl protein, analogous to the recA protein in E. coli, which has been extensively studied (see Section 4 above) Like recA, recl in Ustilago can locate and pair (or synapse) similar DNA sequences. Unlike recA, which requires the presence of topoisomerase (an enzyme which can unwind DNA), recl acts alone.

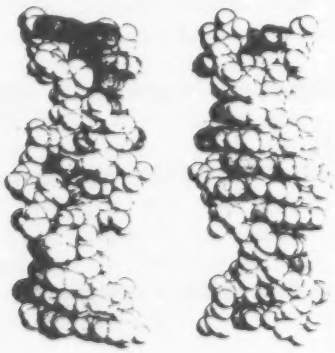
Dr. Holloman used a special system to study pairing and recombination of a double-standed linear DNA molecule (dsDNA) and a single-stranded circular DNA molecule (ssDNA) mixed with recl protein. In this system, recl can attach to the ssDNA and "read" its nucleotide sequence, search the dsDNA for a homologous nucleotide sequence, pair the ssDNA with the dsDNA, and facilitate recombination of the ssDNA into the dsDNA.



To accomplish this, recl must overcome a major difficulty - both the dsDNA and the ssDNA are tightly wound into helices which prevent pairing of sequences. The DNA strands must untwist before recl can match and align the proper sequences. Once the strands are untwisted, the dsDNA and ssDNA can synapse to form a transient structure, called a paranemic joint. An alternative structure, a plectonemic joint, would involve winding all three strands together. The difficulties of creating a plectonemic joint make it an unlikely candidate for the structure of DNA at early synapsis, although it may arise at other times during recombination.

Dr. Holloman knew from the work of Alexander Rich that DNA is found in the cell in one of two forms - normal B-DNA and rare Z-DNA. B-DNA is the common right-handed helix; Z-DNA winds in the opposite, left-handed direction, as shown below:

Z-DNA



B-DNA

Z-DNA is most common in DNA with a high content of the nucleotides cytosine and guanine and is less tightly wound than B-DNA. No clear physiological role of Z-DNA has yet been identified, and Dr. Holloman began to wonder if Z-DNA might be involved in recombination. The fact that it is straighter than B-DNA made it a candidate for the DNA structure found in paranemic joints.

Z-DNA can be made easily in the laboratory. DNA chains made only of the bases guanine and cytosine, called poly dG-dC, are easily converted to the Z conformation by a variety of agents, including concentrated salt or bromine solutions. Bromine is particularly useful because it stabilizes Z-DNA so that it does not flip back into the B conformation when conditions change. Bromine's mode of action is still a mystery; it is possible that bromination is similar to the better-studied process of methylation, which is known to affect many DNA functions, such as transcription and replication.

Dr. Holloman's primary interest was the possible role of Z-DNA in synapsis. He hypothesized that Z-DNA might be found in paranemic joints and soon found that antibodies for Z-DNA bind to paranemic joint molecules and that the appearance or disappearance of Z-DNA (or a Z-DNA-like molecule) parallels the formation and decay of paranemic joint molecules. Dr. Holloman next began a series of experiments to define the relationship of recl and Z-DNA. He began by investigating the abilities of recl to bind to B-DNA and Z-DNA. The assay for binding of recl protein to DNA is simple:

1) Recl protein binds readily to a nitrocellulose filter.

2) DNA passed over the recl/filter complex binds to the recl protein,

3) The amount of DNA retained by the filter is a measure of recl-DNA binding.

Dr. Holloman found that Z-DNA binds to recl more efficiently than does B-DNA.

Another set of experiments utilized the plasmid pLP32, which has 32 alternating cytosine and guanine bases. In the presence of topoisomerase I (an enzyme which affects the amount of twisting in the helix) and ethidium bromide, the poly dG-dC DNA segment flips into the left-handed form. pLP32 was derived from the commonly-used pBR322 plasmid and differs from it only by the special 32-base sequence. In a series of experiments designed to test recl/Z-DNA binding, Dr. Holloman discovered the following:

1) pLP32 in the Z-DNA form binds recl protein more efficiently than does pBR322,

2) pLP32 in the Z-DNA form binds recl more efficiently than does pLP32 in the B-DNA form (that is, not exposed to topoisomerase I and ethidium bromide),

3) recl protein competes efficiently with Z-DNA antibodies for binding to the Z-form of plP32.

4) A brominated strand of DNA containing only guanine and cytosine and in the Z-DNA form competes efficiently with pLP322 for binding to recl protein .

These results, taken with previous work demonstrating that recl promotes recombination, support Dr. Holloman's hypothesis that the formation of Z-DNA, possibly triggered by methylation of special sites, facilitates synapsis of recombining DNA strands. Once Z-DNA is formed, recombination proteins, such as recl, could align DNA strands into a paranemic joint. This view is receiving more attention as supporting evidence accumulates, and it promises to contribute significantly to our understanding of recombination. Uncovering the details of recombination is especially exciting and gratifying because recombination is the basis for generation of variation among individual organisms and is implicated as being vital to many cell functions. The fact that few disorders are known to involve defects in recombination testifies both to the importance of the process (that is, any changes in the recombination pathway are lethal) and to how little we now know about the process. Dr. Holloman's research may shed some light on an important part of the picture.

6. "An Assay for Mismatch Repair"
R01 GM 23719 (Modrich, P.), Duke University

Mutation is a biologically important process, providing the genetic variation necessary for adaptation and evolution. However, most mutations are harmful, leading to dysfunctional or nonfunctional proteins or to changes in gene regulation. Thus, repair of mutations to preserve the integrity of a genetic system is also important. Indeed, several genetic disorders, such as xeroderma pigmentosum, are related to defects in genes that control repair.

Mutations change the sequence of nucleotide bases along a DNA molecule, thereby altering the genetic code. Causes include extrinsic factors, such as uv light and chemical mutagens, as well as intrinsic processes, such as DNA replication and recombination. The integrity of the genetic code depends on accurate replication of the DNA molecule. DNA is a double-stranded molecule, whose two sugar-phosphate backbones are held together by hydrogen bonds between pairs of nucleotide bases. Normally, adenine (A) pairs with thymine (T), and cytosine (C) pairs with guanine (G). During normal replication, the DNA strands separate, and complementary daughter strands are synthesized by matching new nucleotide bases to the sequence of bases on the two parental strands. Two processes can interfere with this accurate duplication of the DNA molecule. First, during normal replication, incorrect bases may be inserted into the new strand creating mismatched base pairs (for example, a G might be paired with a T). If not corrected, mismatches change the genetic code, often leading to changes in protein structure or to alterations in control elements on the chromosomes. Second, just before replication, homologous chromosomes may exchange genetic material in a process called recombination. This usually creates mismatched regions, called heteroduplexes, on the chromosomes involved in the exchange.

Most biological systems have evolved repair pathways to correct mismatches generated during normal DNA replication or recombination. Mismatch repair also appears to be related to other important genetic occurrences such as gene conversion, transformation, and map expansion.

Mismatch repair is carried out by a battery of specialized enzymes which can recognize the mismatch, excise the incorrect base, and replace it with the correct nucleotide. But how does the repair complex recognize which base in a mismatch is incorrect and thus a candidate for excision? To function properly, the complex must leave the parental strand intact while correcting the mutated base in the daughter strand. Radding (GM30488, GM33504) suggested in 1980 that methylation might provide the basis for strand discrimination because newly-synthesized strands are unmethylated. Several lines of evidence support this hypothesis for E. coli. For example; the dam gene product, a methylase, attaches methyl groups to newly-syntheszed DNA strands at sites containing the nucleotide sequence G-A-T-C. Mutants deficient in the dam methylase are unable to repair mismatches properly.

These suggestive studies have prompted much research into the role of methylation of DNA and into how the repair complex works. However, progress has been slow because no simple assay for repair has been available. A simple example illustrates the problem. The replication complex is remarkably faithful, generating mismatches in only 1 in 10^6 to 10^8 bases. An E. coli chromosome contains 4 million bases, so one could expect only four errors for each cell cycle. Testing cells for such infrequent events has been both difficult and imprecise.

Dr. Paul Modrich has recently combined his genetic and biochemical expertise to develop an in vitro assay for mismatch repair. As a biochemist, he was interested in studying the enzymes used for repair but recalled the admonition of one of his professors--"Never look for a protein unless you have a simple assay"! As a geneticist, Dr. Modrich was impressed with the elegance of molecular genetics techniques. The system that grew out of Dr. Modrich's studies is based on the observation that recognition sites for EcoRI restriction endonuclease are resistant to cleavage when they contain mismatches. Dr. Modrich, utilizing a method developed by Drs. Dan Nathans and David Shortle (GM 30649) prepared bacteriophage whose circular DNA contains one EcoRI site with a G-T mismatch. After exposure to the mismatch repair system, repaired molecules are susceptible to cleavage by the endonuclease, but unrepaired molecules will remain circular. Thus, Dr. Modrich can assay for mismatch repair by growing the phage in the E. coli strain of interest and then assaying for linear vs circular phage DNA strands. In addition, Dr. Modrich was able to control the state of strand methylation by propogating the bacteriophage in E. coli hosts with varying amounts of dam gene activity.

Dr. Modrich screened a variety of E. coli mutants to test the accuracy of his assay. MutH, mutL, mutS, and uvrE mutants have high mutation rates and are defective in mismatch repair. After the assay confirmed these findings, Dr. Modrich was able to begin assays for repair enzymes.

Sidney Kushner (GM 27997) has identified the uvrE gene product to be a helicase, one of the topoisomerases responsible for the topological structure of DNA. Dr. Modrich has now isolated and is beginning to characterize gene products of mutH, mutL, and mutS. He has already shown that single-strand binding protein produced by the ssb gene is also required for mismatch correction in the in vitro system. Thus, the assay has proved of great importance. For the first time, scientists can determine the efficiency of a mismatch repair systems and can isolate and purify the proteins involved.

Results of studies on methylation proved very interesting. When both strands or neither strand were methylated, mismatch repair was very inefficient--only 30 percent of its normal value. When only one strand was methylated, the efficiency of the repair complex was high. These results support the hypothesis that mismatch repair enzymes require one methylated DNA strand in order to function properly.

Previous work in Dr. Modrich's laboratory showed that mismatch repair is accompanied by DNA synthesis. The in vitro assay system has allowed Dr. Modrich to determine that this synthesis is localized to specific regions of the DNA, presumably to the regions where nucleotide bases are being replaced. This evidence suggests that mismatch repair involves excision of several bases, not just the incorrect base.

These findings--that mismatch repair requires both localized DNA synthesis and methylation of one DNA strand--is surprising because the G-A-T-C site where methylation occurs is 1,008 base pairs away from the mismatch! This provocative finding has led Dr. Modrich to propose that the repair complex binds to DNA at the mismatched site and, while remaining bound there, processes the DNA until it finds the methylated G-A-T-C sequence. The new strand is then digested and resynthesized between the G-A-T-C site and the mismatch. This allows the repair complex simultaneously to identify parental strand and to correct a mismatch.

Dr. Modrich's assay has stimulated much research in repair systems and has encouraged scientists to develop some specific, testable models of how mismatch repair occurs. Future work will focus on identifying the enzymes in the repair complex and on elucidating their precise mechanisms of action. From these studies, scientists hope to gain a better understanding of how biological systems respond to and correct genetic damage.

7. "Directed Homologous Recombination in Human DNA, and Its Relevance to Gene Therapy"
   R01 GM 20069 (Smithies, O.), University of Wisconsin
   R01 GM 33943 (Kucherlapati, R.), University of Illinois

Over 3,000 human disorders caused by defects in single genes have been identified, and at least 1 out of every 100 babies born in the U.S. is affected by one of these genetic disorders. In developed countries, nearly 50 percent of all infant deaths have genetic causes. Despite the rapid advances in our knowledge of genetics and the development of technologies to identify and study genetic disorders, there are no cures for any genetic disorders. Medical support is currently limited to treatment of the symptoms of only a few of the most common disorders.

There is hope that the situation will improve in the future. Drugs made by recombinant DNA techniques are currently in use; a few genetic disorders can be detected and even treated prenatally; and our understanding of the genetic bases of many disorders has been greatly advanced by scientists using state-of-the-art genetic technologies. It is not surprising that many researchers are now trying to utilize the concepts and technologies of modern genetics to attempt to cure genetic disorders.

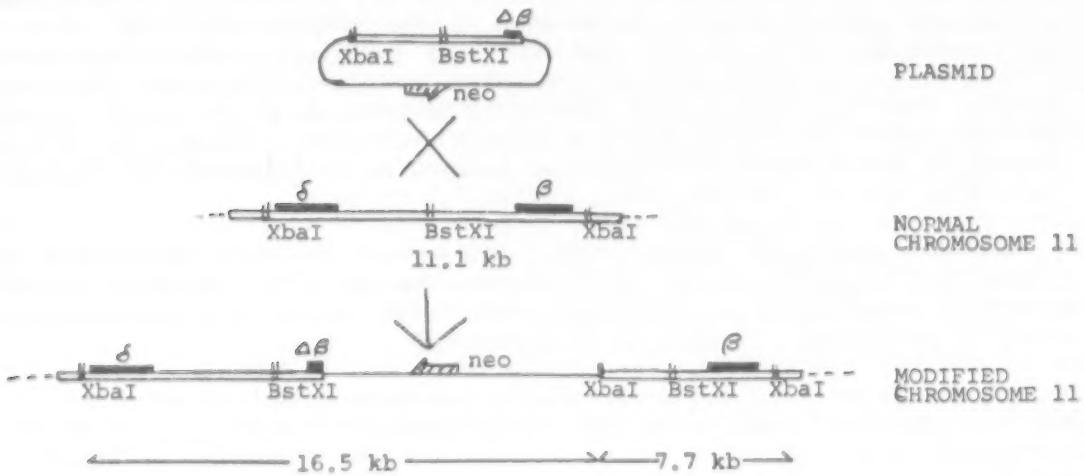
Curing a genetic disorder involves replacing the defective gene with a normal one in some or all of the body cells. There are many variants of gene replacement therapy, some of which are currently being studied in animal models. Some may be applied to human disorders within a few years. The most likely procedure would involve transplants of genetically modified bone marrow. Many disorders, such as the thalassemias, are caused by defects in genes which are expressed in the bone marrow stem cells. The symptoms of other disorders, such as adenosine deaminase deficiency or Lesch-Nyhan syndrome, might be relieved if normal genes could be expressed in bone marrow, thereby releasing needed enzymes into the circulation. In some cases, genetic disorders have been successfully treated by replacement of some of the patient's defective marrow with normal marrow from an immunologically compatible individual. The most common reason for failure of this procedure is rejection of the transplanted tissue by the patient's own immune system. This problem could be avoided by using the patient's own bone marrow, into which a normal gene had been inserted. Although it is an important first step, this procedure is not without problems. One of the most important problems is that the inserted gene may randomly integrate into the patient's own chromosomes, disrupting normal genes or regulatory regions. Uncontrolled or inapproporiately regulated expression of normal genes could be as serious as the orginal disorder.

A better solution would involve replacing the defective genes with normal ones. Safe and reliable gene replacement therapy requires that the inserted, normal gene pair and recombine with the defective gene, specifically replacing only that gene. This process--called site-specific recombination--resembles the naturally occurring process of homologous recombination. Since T.H. Morgan's experiments in the 1910's, scientists have known that genes on homologous chromosomes can recombine by exchanging DNA strands, as in crossing-over. Sequences on DNA fragments may also integrate into regions of chromosomes where the DNA sequences are very similar. Homologous recombination often occurs naturally but is not understood in any detail. For example, we know that lambda phage recombine into homologous regions in the E. coli chromosome; plasmids do not normally integrate into the bacterial chromosome; and Mu phages integrate nonspecifically (i.e. at many nonhomologous sites). The reasons for these differences in recombination styles are unknown but suggest that competing pathways for recombination in bacteria may exist. Understanding and eventually controlling homologous recombination is an important step toward replacing defective genes with normal ones.

In 1977, Richard Axel developed a system, called calcium phosphate transfection, for inserting fragments of DNA into eukaryotic cells. This procedure made it possible for scientists to study homologous recombination in cells of higher organisms. Initial results were as complex as those from the bacterial system. Homologous recombination occurred in fairly high frequency in yeast; however, in mammalian cells, the frequency appeared to be very low. Recently, several researchers--among them, Mario Capecchi (GM21168)--have begun to define the conditions needed to obtain homologous recombination in mammalian systems. Their studies have shown that in carefully designed and controlled experiments, high rates of homologous recombination are possible.

Building on this work, Drs. Oliver Smithies and Raju Kucherlapati have recently completed a set of experiments which can justifiably be called a significant breakthrough in gene replacement therapy studies. They have inserted a beta-globin gene into a naturally-occurring, predetermined site in a human chromosome.

The system is complex, but several important features stand out and are illustrated in the figure below.



1. The recipient cells were mouse/human hybrids, in which most human DNA had been eliminated. The only human DNA in the cells was an X chromosome onto which a portion of chromosome 11 had been attached. The target sequence, contained in the globin gene region (both alpha and beta genes are present), was on the chromosome 11 portion of the chromosome. These cells express the human adult beta-globin chain at high levels.

2. A new technique, called electroporation, was used to insert the vectors, such as plasmids, into the cells. This procedure, developed largely by Drs. Huntington Potter and David Dressler (GM17088) uses an electric shock to introduce transient holes in the cell membrane. Vectors can be taken up in this brief time. The technique is much more reliable than other procedures for inserting one, and only one, vector into each recipient cell.

3. The beta-globin gene to be inserted into the chromosome was incomplete and, if expressed, would produce a shortened globin chain. While the short chain might not be functional, it could serve as a marker to determine if the inserted genes were expressed.

4. The plasmid contains a gene (neo) which confers neomycin resistance to the cell. Normally, cells are killed by exposure to this drug, but cells which had integrated the plasmid were resistant. This enabled Drs. Smithies and Kucherlapati to detect the total number of integrations, both specific (at the target sequence) and nonspecific (at other sequences).

5. The beta sequences on the plasmid and the chromosomal DNA provided a region of homology which directed the plasmid to the target site. Cutting the plasmid at the BstXI restriction site before inserting it into the cells facilitated recombination at the BstXI site.

6. The BstXI target sequence is flanked by two XbaI restriction enzyme sites. The XbaI sites are 11.1 kb apart in the normal chromosome. Therefore, treatment of the normal DNA with XbaI restriction enzyme yields an 11.1 kb fragment, detectable by gel electrophoresis. However, when the plasmid, which contains another XbaI site, is inserted at the target site, the composite chromosome contains three XbaI sites. Integration of the plasmid at the correct site could be assayed by the presence of two XbaI fragments, one 16.5 kb long and the other 7.7 kb long.

7. The frequency of site-specific integration could be calculated by comparing the number of cells containing the two XbaI fragments (site-specific integrations) to the number of cells which were converted to neomycin resistance (nonspecific integration).

The experiments were complicated and lengthy but definitive. As expected, the plasmid was capable of integrating into the chromosomes at many sites; however, the frequency of site-specific integration was far above that expected by chance. Site-specific integrations occurred with a frequency between 1/200 and 1/1000. Only one copy of the plasmid was integrated in each cell. Further, it appears that the inserted DNA sequences were capable of being expressed; however, definitive studies to determine if the shortened beta-globin chain was made are not yet complete.

These experiments establish that a planned modification of a specific human gene can be accomplished without disrupting the remainder of the genome and can preserve the ability of the integrated genes to be expressed. They foster the hope that specific, planned modifications of the beta-globin locus in bone marrow cells will be feasible as a treatment of patients with genetic disorders such as thalassemia and sickle cell anemia. This work is clearly a breakthrough in achieving that distant goal.

D. Extranuclear Inheritance

1. Overview

Not all of the genetic information of an organism is localized to chromosomes of the nucleus. DNA is also found in some cytoplasmic organelles, notably in mitochondria, the cell's centers for energy production, and in plant chloroplasts, which carry out photosynthesis. These are organelles which grow and divide, and whose DNA replicates in a manner similar to that of chromosomal DNA.

Female gametes, such as the mammalian egg and plant ovum, are rich in cytoplasm and cytoplasmic organelles. Sperm or pollen contain little or no cytoplasm. Thus, at the time of fertilization, the male gamete usually contributes few if any mitochondria or chloroplasts to the new organism. Thus, since genetic traits carried on the mitochondrial or chloroplast DNA are transmitted via the maternal gamete, this type of inheritance is sometimes referred to as maternal, or cytoplasmic, inheritance.

However, most of the proteins found in mitochondria and chloroplasts are specified by nuclear genes. It is still a mystery why a few proteins are specified by the organelles themselves and are made in situ (in place), while most are specified by the nuclear chromosomes, made in the cytoplasm, and transported to specific sites in organelles. The mechanism of this transport from cytoplasm to organelle is the topic of Section Three of this group of highlights. It turns out that there are "leader peptides" which target proteins to sites on mitochondria and chloroplasts. Higher plants serve as particularly useful models for studying targeting of proteins to organelles.

Cashmore has found that he can engineer genes and transform plants to produce new proteins in response to regulatory signals such as light. Similar to the work in animal systems and mitochondria, he can target these proteins to chloroplasts if the transforming DNA contains sequences specifying leader peptides.

In summary, much work is now underway to seek greater understanding of many facets of extranuclear inheritance. It is highly likely that some human genetic diseases will be traced to defects in mitochondrial DNA, or defects of mitochondrial protein transport. Although not described below, many population geneticists are beginning to study variants of mitochondrial DNA as tools for understanding genetic diversity within and among populations. Furthermore, work on plants has demonstrated that a thorough understanding of the relationship of the nuclear and extranuclear genes is required to engineer variations in photosynthetic organelles. This work with plants has obvious implications for similar manipulations of animal cells.

2. "Identification of Six New Genes in the Human Mitochondrial Genome"
ROl 11726 (Attardi, Guiseppi), California Institute of Technology

Mitochondria, the organelles responsible for cellular respiration, contain their own DNA. In mammals that have been studied, the mitochondrial DNA (mtDNA) is in the form of circles consisting of approximately 16,000 base pairs. The complete mouse, bovine, and human mtDNA sequences have been determined and were found to be similar in their gene organization. Each mammalian mtDNA molecule contains genes for each of the two mitochondrial ribosomal RNAs and for 22 mitochondrial transfer RNAs. There are also 13 genes which, based on their sequences, are thought to code for proteins. The proteins encoded by five of these genes were deduced from their homology to yeast mitochondrial genes which had been identified by biochemical and genetic analyses. Identification of these genes in mammalian mtDNA was subsequently confirmed by the reaction of specific antibodies with the products of these genes. The structural genes which have been identified in this way include cytochrome oxidase subunits I, II, and III, ATPase subunit 6, and apocytochrome b.

The remaining eight putative protein-coding genes have no obvious homology to yeast mitochondrial genes. They are known as unidentified reading frames (URFs). Reading frames are DNA sequences that contain appropriate start and stop signals and have three-nucleotide codons that correctly specify amino acids. Most of the URFs are known to be transcribed into RNA. The proteins they encode would be very hydrophobic, most likely enzyme complexes located in the inner membrane of mitochondria.

A group of investigators, including Guiseppi Attardi, have recently identified the products of six of the remaining URFs from human mtDNA. This advance will add substantially to our understanding of mitochondrial biogenesis and has interesting implications for the evolution of the mitochondrial genome.

Dr. Attardi's group speculated that, among the enzyme complexes of the inner mitochondrial membrane, the NADH dehydrogenase complex (also called NADH-ubiquinone oxidoreductase) seemed to be a good candidate for containing some of the URF products. This complex, which is part of the respiratory chain, is made up of a shell of approximately 15 hydrophobic polypeptides surrounding a core containing the NADH dehydrogenase enzyme. Based on several lines of evidence, Dr. Attardi and his colleagues have demonstrated that six URFs encode six components of the NADH dehydrogenase complex.

First, they tested the idea that the URF products are in fact part of a protein complex. To do this, they obtained peptides synthesized by several of the URF genes and raised antisera against them. Each antiserum precipitated not only the peptide against which it was raised, but also the products of the other URFs. This suggested that the URF products are associated with each other in some sort of complex.

Next, they incubated cells with radioactive protein precursors under conditions where cytoplasmic protein synthesis is blocked but mitochondrial protein synthesis proceeds. The mitochondria were then isolated and solubilized to release the membrane proteins. The proteins were allowed to react with antiserum raised against a highly purified preparation of the native bovine NADH dehydrogenase complex. The antiserum reacted specifically with the products of six URFs, causing each to precipitate. The precipitated URF products which had incorporated radioactive label could be detected after electrophoresis. The antiserum did not react with other mtDNA gene products, including cytochrome oxidase subunits I, II, and III. And as expected, antiserum against the cytochrome oxidase subunits did not react with the URF gene products. These results demonstrated that the NADH dehydrogenase complex contains polypeptides encoded by six of the mtDNA URFs.

As further evidence that the URFs encode polypeptides in the NADH dehydrogenase complex, Dr. Attardi and his colleagues subjected the newly synthesized URF products to a fractionation scheme designed to separate various mitochondrial membrane proteins. Detectable levels of NADH dehydrogenase enzyme activity correlated only with those fractions that contained URF products.

The molecular weights of the URF products shown to be a part of the NADH dehydrogenase complex are 51,000 daltons, 36,000-39,000 daltons, 25,000 daltons, 24,000 daltons, 6,000 daltons, and 3,500 daltons. These are in reasonable agreement with the molecular weights reported for some of the proteins in the bovine NADH dehydrogenase complex. Dr. Attardi's group believes that at least some of the URF products are part of the shell of hydrophobic proteins surrounding the catalytic subunits.

It is of interest that yeast mtDNA does not contain any genes homologous to the mammalian mtDNA URFs that encode components of the NADH dehydrogenase complex. The absence of these genes in yeast may relate to the structural and functional differences that have been observed in the NADH dehydrogenase region of the respiratory chain in yeast compared to that of mammals. Alternatively, it is possible that in yeast some or all of the NADH dehydrogenase polypeptides are encoded in the nucleus and are transported to the cytoplasm. The finding of variability in the genetic control of NADH dehydrogenase in different organisms extends similar observations previously made on another mitochondrial enzyme complex, H^+-ATPase complex. It also suggests that mitochondrial evolution might have followed very different pathways in different organisms. In some organisms, genes encoding mitochondrial proteins are located in the mitochondrial genome, while in others, the homologous genes are located in the nuclear genome. The significance of such partitioning remains to be determined.

3. "Promising Leads in Understanding Leader Sequences of Mitochondrial Proteins"
P50 GM 32156 (Rosenberg, L.), R01 GM 34433 (Horwich, A.), R01 GM 29765, R01 GM 33904 (Mellman, I.), Yale University

Although mitochondria have their own DNA and protein synthesizing machinery, nearly all of the several hundred proteins contained in mitochondria are encoded by nuclear DNA and are synthesized in the cytoplasm. A question of current major interest is how these proteins find their way to the mitochondria and, once there, how they end up in the correct part of this complex organelle. Targeting of mitochondrial proteins is unlike that of proteins destined for lysosomes, cellular membranes, or extracellular transport. The latter are synthesized on membrane-bound polyribosomes and are inserted into the cisternae of the endoplasmic reticulum for transport. In contrast, nuclear-encoded mitochondrial proteins are synthesized on free ribosomes and are subsequently released into the cytoplasm for import by the mitochondria. Understanding how proteins are targeted to mitochondria is of more than academic interest, since there is a suspicion that certain inborn errors of metabolism may be due to defects in protein targeting to mitochondria.

For many years Dr. Leon Rosenberg has been studying various aspects of the synthesis, regulation, transport, and processing of the protein ornithine transcarbamylase (OTC), which is a liver mitochondrial enzyme involved in the synthesis of urea and the detoxification of ammonia. Patients without functional OTC often die in infancy from a buildup of ammonia. Some of these patients do not synthesize OTC, others make non-functional OTC, and still others probably make OTC that does not get into their mitochondria.

Dr. Rosenberg, in collaboration with Dr. Arthur Horwich, a former postdoctoral trainee (T32 GM 07439) in his laboratory, and Dr. Ira Mellman, have reported significant progress in understanding how OTC, and probably other mitochondrial proteins, are targeted. (Drs. Horwich and Mellman are supported by the Cellular and Molecular Basis of Disease Program.) It was shown previously that OTC is composed of three identical subunits, which are synthesized as larger precursors. Import of the precursors involves specific binding by receptor molecules present in the outer mitochondrial membrane and translocation through the mitochondrial membranes by an energy-requiring process. Inside the mitochondria, enzymatic removal of a 32-amino acid leader sequence at the amino-terminal end of the subunits and assembly of the three mature subunits produce an active enzyme. It was also shown that the precursor containing the leader sequence is taken up by isolated mitochondria, but the corresponding mature protein, lacking the leader sequence, is not taken up. It therefore seems certain that the leader sequence is required for recognition.

Drs. Rosenberg and Horwich and members of their laboratories compared the leader sequence of human OTC with the leader sequences of mitochondrial enzymes from yeast, Neurospora, and rat. Although the leader sequences do not have any sequence homology, they do share several features in their amino acid composition. The leaders contain no stretches of hydrophobic amino acids, a characteristic of "signal" sequences of secreted proteins. They are devoid of acidic residues which are present in the coding sequence. And they contain an average number of basic amino acids when compared with other eukaryotic proteins.

To define further the characteristics of the leader sequence required for mitochondrial import, Drs. Rosenberg, Horwich, and Mellman and their associates developed a clever cell culture system capable of expressing OTC. Since OTC is only expressed in significant amounts in intact liver, a cell culture system had not been available before. Their strategy involved constructing a plasmid containing the OTC gene joined to the dihydrofolate reductase (DHFR) gene and viral regulatory sequences, transfecting HeLa cells with the hybrid DNA sequence, and selecting for DHFR gene amplification in the presence of methotrexate. As the gene for DHFR became amplified under these conditions, so did the gene for OTC. In cells containing the amplified OTC gene, OTC enzymatic activity was comparable to that in normal human liver, and OTC was found to be localized correctly in mitochondria.

Using this system, they showed that when they substituted the amino acid analog canavanine for the positively charged arginine residues, of which there are four in the leader sequence, OTC was not taken up by the mitochondria. In contrast, when they substituted an analog for the neutral amino acids phenylalanine and methionine, of which there are six in the leader sequence, OTC uptake was normal. These results suggested that the arginine residues in the leader sequence are required either for the import of the OTC precursor into the mitochondria or for the removal of the leader sequence once the precursor was taken up.

Three possible explanations exist for the role of the arginines in the leader sequence. First, the positive charge of arginine itself may be important for recognition. Second, the particular sequences of which the arginines are a part may be required. And third, the conformation of the leader sequence may be changed by an amino acid substitution, making the sequence non-functional.

Experiments in progress using site-directed mutagenesis to engineer OTC precursors with substitutions for specific amino acids in the leader sequence should show which of these explanations is correct.

A long-term goal of Dr. Rosenberg's laboratory is the clinical diagnosis and treatment of OTC defects and other inherited metabolic diseases. He and his colleagues are pursuing this goal by analyzing DNA from patients deficient in OTC to define the molecular nature of specific mutations. The studies described in this Highlight will complement that effort by suggesting possible mutations in the leader sequence that may be associated with OTC deficiencies.

4. "Manipulation of Plant Genetic Systems"
ROl GM 31137 (Cashmore, A.R.), Rockefeller University

A major recent advance in the field of plant molecular genetics has been the successful development of the Agrobacterium-mediated transformation system. Agrobacterium is a soil microorganism which infects wound sites and causes a tumorous disease known as crown gall in many dicotyledonous plants. This infection is an example of naturally occurring gene transfer in that genetic material is passed from the bacterium into the plant where it integrates into the chromosomal DNA. During the last decade a few groups of plant molecular biologists in the United States and Europe have intensively studied the molecular mechanism of certain aspects of this gene transfer. These studies have culminated in the development of procedures whereby it is now possible to introduce probably any gene of choice into plants susceptible to Agrobacterium. Furthermore, modified Agrobacterium vectors have now been developed which still mediate gene transfer but no longer induce the tumorous state. These developments, combined with the fact that single cells from many plant species can be regenerated into plants, set the stage for a new era in the manipulation of plant genetic systems.

Given that we can introduce foreign genes into plants, it is appropriate to ask whether these genes will be expressed and if so is it possible to manipulate their expression. Similarly, we can ask where in the cell will the polypeptides encoded by these genes reside, and is it possible to determine genetically the sites of residence of these polypeptides. Dr. Anthony Cashmore's laboratory has addressed these questions in collaborative research projects with the laboratory of Marc Van Montagu and Jeff Schell in Belgium.

The genes of primary interest in Dr. Cashmore's laboratory are nuclear genes encoding polypeptide components of chloroplast proteins. These polypeptides are the small subunit of ribulose-1,5-bisphosphate carboxylase (rbcS) and the constituent polypeptides (CAB) of the light-harvesting chlorophyll a/b protein complex. The nuclear genes encoding these polypeptides exhibit interesting regulatory characteristics including the fact that they are specifically expressed in leaf tissue and they are induced by light. An experiment was carried out aimed at defining the region of the rbcS gene which conferred light-inducibility. A chimeric gene was constructed in which a promoter fragment from a pea rbcS gene was fused to a bacterial gene encoding chloramphenicol acetyl transferase (CAT). Agrobacterium was used to introduce this chimeric gene into tobacco cells and the resulting transformed callus tissue was grown either in the dark or the light. The chimeric gene was shown to be

expressed in the light-grown tissue but not in the dark-grown tissue. Thus the experiment both defined the region of the rbcS gene responsible for conferring light-inducibility and demonstrated for the first time that it is possible to manipulate the expression characteristics of foreign genes in plant cells.

Dr. Cashmore's laboratory has extended these studies to further define the nature of the rbcS regulatory sequences responsible for light-inducibility. It was shown that these regulatory sequences have characteristics similar to the enhancer sequences which regulate the expression of certain animal cell genes including immunoglobulin and steroid-induced genes. Of particular interest was the demonstration that by fusion to a heterologous promoter, the rbcS regulatory sequences confer the ability to promote light-inducibility. Specifically, the regulatory sequences from the rbcS promoter were fused to the promoter for the nopaline synthase gene. The nopaline synthase promoter normally promotes constitutive expression. However, the fusion promoter containing the rbcS regulatory sequences was shown to have acquired the ability to promote light-inducibility in a manner not significantly different from the rbcS promoter. The significance for plant genetic engineering of this demonstration is that one now has the ability to confer additional regulatory characteristics to genes by using these enhancer-like sequences. Almost certainly other plant genes will be regulated by similar enhancer-like sequences, and these also will prove useful tools in the future manipulation of plant genetic systems.

With the use of "disarmed" Agrobacterium vectors it is possible to introduce genes into plant cells without inducing tumors, and with certain plant cells it is possible to regenerate plants from these transformants. In these regenerated plants one can look for tissue-specific expression of introduced genes. By these procedures a chimeric gene containing a pea CAB gene promoter joined to a bacterial gene, neomycin phosphotransferase (NPT), was introduced into tobacco plants. Tissue specific expression of this bacterial gene was found. The chimeric gene is expressed strongly in leaf tissue and not in root tissue. Thus, in addition to being able to confer light-regulated expression to foreign genes, this experiment demonstrates the ability to confer tissue specific expression.

Both the rbcS and the CAB polypeptides are synthesized outside the chloroplast on cytoplasmic ribosomes. The primary products of synthesis are precursor polypeptides characterized by amino terminal extensions referred to as transit peptides. These transit peptides (analogous to leader sequences of precursors of mitochondrial proteins, See D 3) mediate the import of the precursors into chloroplasts. It was of interest to determine if these transit peptides could be utilized to mediate the import of foreign polypeptide into chloroplasts. Dr. Cashmore's laboratory constructed a chimeric gene in which the rbcS promoter and the coding region for the transit peptide were fused to the bacterial NPT gene. This chimeric gene encodes a polypeptide in which the rbcS transit peptide is fused to the amino terminus of the NPT protein. This gene was introduced into tobacco cells, and transformed photosynthetic tissue was generated. They then isolated chloroplasts from this tissue and assayed for activity. These chloroplasts contained NPT activity in contrast to chloroplasts from tissue transformed with an NPT chimeric gene lacking the rbcS transit peptide. The experiment clearly demonstrated that it is possible to manipulate genetically the transport characteristics of foreign proteins in plant cells.

The experiment concerned the bacterial protein NPT; however it should be possible to manipulate in a similar way the import into chloroplasts of agriculturally important proteins such as those conferring resistance to herbicides. Such experiments are in progress in Dr. Cashmore's laboratory.

In conclusion, it was shown that it is possible to manipulate the expression characteristics of foreign genes in plants. It was demonstrated that photoregulation of plant genes is mediated by sequences similar to animal enhancer sequences. These regulatory sequences, like the animal enhancer sequences, have the ability to confer regulation via heterologous promoters. Probably other plant genes are also regulated by enhancer-like sequences. These sequences will prove to be powerful tools in conferring additional regulatory characteristics to genes under the control of plant promoters. Finally, the Cashmore laboratory has shown that it is possible to engineer genetically the transport characteristics of polypeptides in plants.

E. Control of Gene Expression and RNA Splicing

1. Overview:

Much progress has been made in the past several years in understanding the mechanisms that regulate the expression of genes. In this section, two major areas of research are discussed. The first relates to the tremendous accumulation of information about RNA splicing, a process by which portions of a newly synthesized RNA molecule are selectively removed. Although molecular biologists have identified several distinct types of RNA splicing reaction, it is likely that they all in some way regulate gene expression.

The second area involves research on the proteins and other factors which turn genes on and off. Recent work on Xenopus shows that several proteins form a complex which regulates gene activity by binding to sites on the DNA. In several viral and bacterial systems, investigators have described by crystallography the interaction of proteins with DNA at the atomic level. Of great interest is the common helical structure shared by all the DNa binding proteins so far crystallized.

2. "RNA Splicing: Three Themes with Variations"

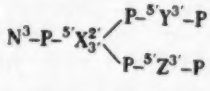
ROl GM 28039 (Cech), University of Colorado; ROl GM 33977 (Green), Harvard University; ROl GM 32585 (Edmonds), University of Pittsburgh; ROl GM 32467 (Sharp), Massachusetts Institute of Technology; ROl GM 32511 (Peebles), University of Pittsburgh; ROl GM 30579 (Matthews), Cold Spring Harbor Laboratory

Some of the most exciting activity in molecular biology today is occurring in the field of RNA processing. Much of this work has been presented, discussed, and coordinated at the three RNA Processing Meetings at Cold Spring Harbor which were sponsored by NIGMS. At the second of these meetings, in 1984, Dr. Cech presented the evidence for "autocatalysis" of RNA, a discovery which has changed our view of enzymology as well as RNA processing. This work also demonstrated the careful, elegant manner with which RNA processing is being examined by investigators in this field and the fundamental importance of their findings.

The extent and significance of RNA processing was appreciated with the recognition, about 5 years ago, that the coding regions of many eucaryotic genes are interrupted by stretches of noncoding DNA called intervening sequences (IVS), or introns. These introns are transcribed as part of precursor RNAs and are subsequently deleted by a cleavage-ligation process termed "RNA splicing". It is premature to believe we have a final count of the ways in which splicing is achieved. However, on the basis of findings of the past three years, it now appears that there may be three general classes of RNA splicing reactions—one for nuclear-encoded mRNAs, a second for tRNAs, and a third for nuclear-encoded rRNAs and mitochondrial-encoded mRNAs and rRNAs.

These three classes of RNA splicing reaction can be readily distinguished on the basis of how the splice sites (the phosphodiester bonds to be cleaved and ligated) are specified. A summary diagram of the reactions is provided below. An appendix is attached for those who enjoy looking at more detail.

In addition to the fundamental discovery that certain RNAs can function as catalysts, another important discovery has been that many, if not all, introns are excised as circular molecules, in some cases referred to as "lariats". This discovery resolved a "mystery" which had puzzled investigators for some time. Several years ago, a graduate student of Dr. Mary Edmonds (GM 32585) had noted that a species of RNA found in the nucleus behaved in a fashion which could best be explained by a branched structure. The chemical structure of this branch was published in 1983:

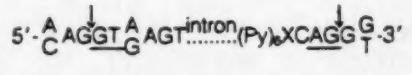


These branched structures are now known to be at a node connecting the circular part of the "lariat" to its linear portion. The discovery of branched and circular structures has altered our view of the chemistry and function of RNAs in cells in a profound way.

The three forms of splicing are compared and contrasted below.

(1) Nuclear mRNA Precursors

For nuclear-encoded mRNA precursors, the splice sites always contain a sequence resembling



The underlined residues are invariant, whereas the remainder of the conserved sequence displays considerable variation.

It has been shown in the gene for β-globin that this pair of conserved sequences, flanking the intron on both sides, includes the site of the RNA splices. A variety of genetic experiments, including base changes, insertions and deletions in various portions of the exons and within the introns have pinpointed the sequences which result in splicing.

It has been postulated that small nuclear ribonucleoproteins (snRNPs) recognize these highly conserved sequences. It has been found that splicing of some adenovirus mRNAs is inhibited by anti-U1-snRNP antibodies. In addition, there is now direct physical evidence of selective binding of the U1-snRNP to the 5' splice site of an artificial β-globin pre-mRNA.

It has also been postulated that heterogeneous nuclear RNA (hnRNA) or, when complexed with protein, heterogeneous nuclear ribonucleoprotein (hnRNP) may bring the 5' and 3' splice sites near each other, and thus allow splicing to proceed. This idea may serve to explain why certain splice sites (cryptic splice sites) are not normally utilized despite having the requisite sequence, except when the usual splice sites are inactivated by mutation. Normally, the usual pair of splice sites are brought into juxtaposition, aided by hnRNA or hnRNP. When the mutation precludes this, the cryptic splice sites, assumed to be more buried in the three-dimensional mRNA structure and thus kinetically less favored, are utilized instead.

Until very recently, little information was available regarding the biochemical mechanism of nuclear pre-mRNA splicing. Within the last year, however, several research groups have had success in reproducing this finicky reaction in vitro. Michael Green has examined the excised intron from human β-globin mRNA precursor splicing reactions. These introns exhibit several unusual properties: their electrophoretic mobility is anomalous, primer extension is blocked at the position of the RNA branch, and there is a nuclease-resistant component. In vitro, the excised IVS, which is 143 nucleotides long, is slowly converted to a 130 nucleotide RNA product which still exhibits anomalous electrophoretic mobility. Both products are in the form of "lariats".

Dr. Green has detected and partially characterized a novel RNA processing enzyme that converts these RNA lariats to linear molecules by specific cleavage of the 2',5'-phosphodiester bond (see diagram, A). The activity of the debranching enzyme is very specific: the 129 normal 3',5'-phosphodiester bonds in the 130 nucleotide RNA are left intact while the single 2',5'-phosphodiester bond is efficiently cleaved. Both the lariat and linear RNA forms are found in vivo. Regardless of its biological role, the enzyme already has proven enormously useful in experiments aimed at characterizing splicing intermediates and products.

(2) tRNA Precursors

Unlike mRNA precursors, tRNA presursors have no conserved sequence near their splice sites. After examining conserved sequences throughout the molecules and after performing genetic experiments with insertions in the exons near the splice sites and in the introns, investigators have concluded that splicing of introns from tRNA precursors is determined by the secondary and tertiary structure of the remainder of the molecule (these sequences also determine the conformation of the mature tRNA). A single yeast splicing endonuclease can excise precisely the intervening sequences from ten different tRNA precursors.

The tRNA precursors have introns of different sizes and sequences, but they have exons with similar, highly conserved secondary structures.

The mechanism of tRNA splicing has been described for yeast, and the enzymes have been partially purified (Peebles, see diagram, B-1). The mechanism is unusual. A 3', 5' phosphodiester-2' phosphomonoester linkage is formed in which the phosphate that forms the splice between the exons is not a phosphate from the original pre-tRNA molecule, but is donated by ATP. In wheat germ and Chlamydomonas, tRNA is spliced in vitro by the same mechanism.

A different situation occurs with tRNA splicing in HeLa cells (Filipowicz and Shatkin) and probably in Xenopus oocytes (see diagram B-2). The phosphate that forms the splice is one that was present in the precursor molecule at the 5' end of the intron (see appendix B-2, 1). There is an ATP requirement for the reaction in HeLa cells. It is not known whether the ATP-dependent cyclic phosphate-forming enzyme ("cyclase") has an obligatory role in the reaction.

In spite of the differences, the yeast/wheat and HeLa/Xenopus splicing mechanisms have several unifying features: splicing is a two-step reaction, cleavage and ligation are catalyzed in separate reactions, the endonuclease cleaves the RNA chain so as to leave the phosphate on the 3' end, and both mechanisms involve 2',3'-cyclic phosphates and probably involve ATP as an energy source.

(3) Nuclear rRNA and Mitochondrial mRNA and rRNA Precursors

In the remaining class of splicing reactions, the intron plays a much larger role in its own excision. In the case of the nuclear rRNA genes of Tetrahymena, splicing of the rRNA precursors occurs by an autocatalytic reaction that, at least in vitro, requires no protein (Cech, see diagram, C). The excised intron is capable of undergoing a second splicing reaction to produce a circular form. The entire cleavage-ligation activity resides within the intron itself.

Tetrahymena rRNA precursor is capable of self-splicing in the absence of proteins, while the splicing of many mitochondrial mRNA introns does require proteins. In yeast mitochondria, splicing is blocked both by mutations in the exon and by mutations in nuclear genes, indicating the involvement of proteins. One model holds that, in mitochondria, proteins are required to help the intron fold into its productive structure, a structure which is self-splicing.

Splicing of rRNA precursors in Tetrahymena takes place by a completely "self-catalyzed" reaction and the chemistry of the cleavage-ligation is different from splicing of tRNA. Splicing occurs by a series of two trans-esterification reactions which convert one phosphate ester to another without any intermediate hydrolysis. Since the phosphodiester bond is conserved throughout, neither ATP nor GTP is required for hydrolysis. A free 3' hydroxyl group is required which is contributed by a guanosine cofactor. The reaction is characterized by the addition of a nucleotide at the 5' end. The IVS is excised as a unique linear molecule. This linear molecule is subsequently converted into a circular molecule by an autocyclization process. During cyclization, a 15-nucleotide segment ("15-mer") is produced which is capable of forming a hairpin structure. The cyclization process is readily reversible. It is not known whether mitochondrial splicing follows the same reaction, but the essential internal sequences are conserved, which suggests the mechanisms will be related.

In one group of yeast mitochondrial mRNA genes, mutations which affect splicing have been found within introns, but far from the splice junctions. Once again, however, the structure of the intron determines splicing. There are four sequences, each 10-12 nucleotides long, which appear to be essential. There is evidence that these four sequences, which are highly conserved, form two base-paired stems, which help form the secondary structure of these introns.

SUMMARY

Based on the way in which splice sites are specified, all split genes can be grouped into three classes: nuclear mRNA splice sites are designated by short sequences at the intron-exon junction; nuclear tRNA splice sites are designated by the three-dimensional structure of the exons; and nuclear rRNA and mitochondrial rRNA and mRNA splicing reactions are designated largely by the structure of the intron. These three classes appear to be distinct with respect to the chemistry of the splicing reactions.

The three classes of intron appear to have evolved independently, so they may be serving quite different functions. However, the discovery that many or all introns are excised in a circular form directs our attention to the possible roles of these circular molecules. The most intriguing possibility arises from the fact that the circular form provides immediate protection of the molecule--no ends are available to be degraded. The possibility that the excised IVS have regulatory functions is being examined.

DIAGRAM

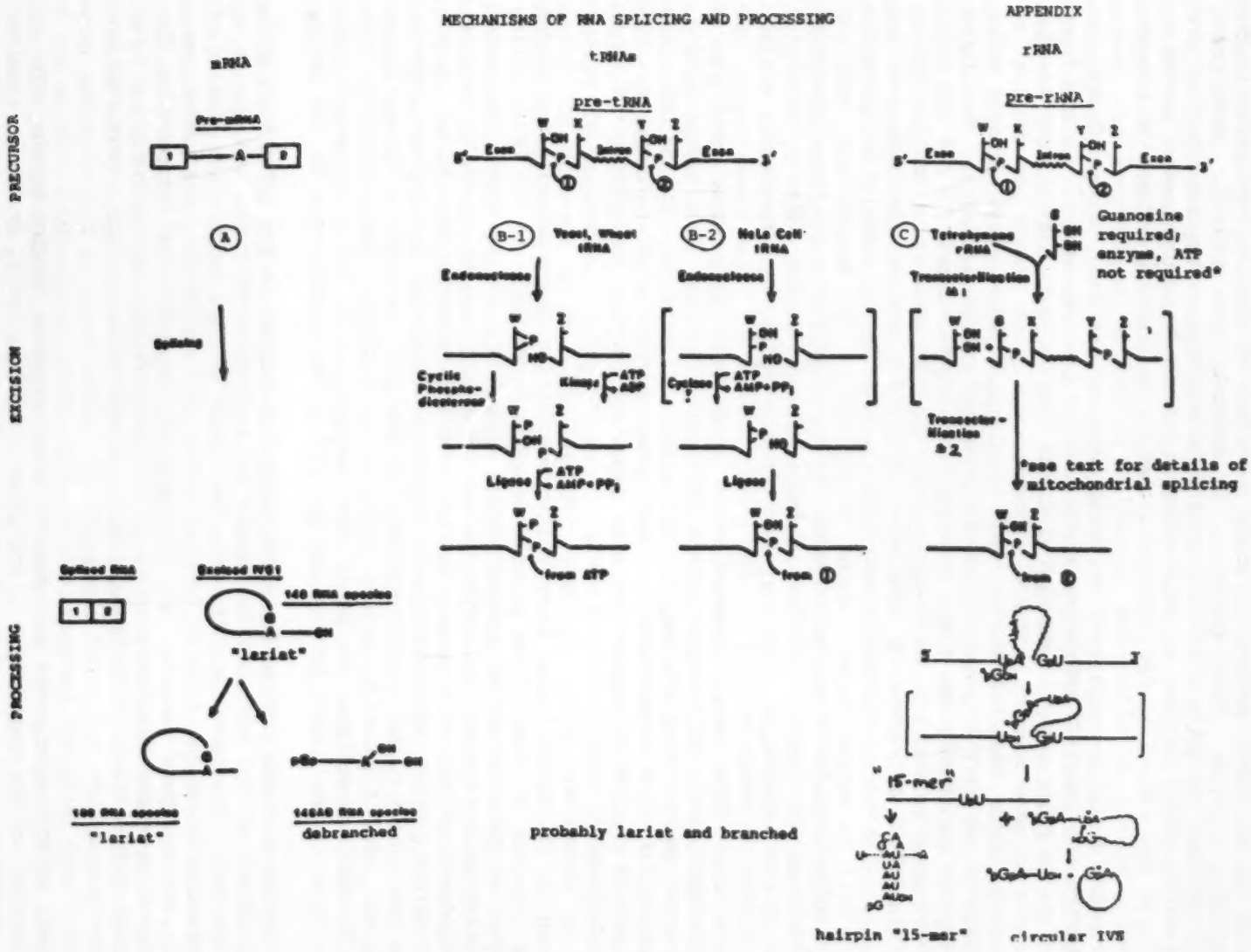
MECHANISMS OF RNA SPLICING AND PROCESSING

	mRNA	tRNAs		rRNA
PRECURSOR:	pre-mRNA (exon-intron-exon)	pre-tRNAs (exon-intron-exon)		pre-rRNA (exon-intron-exon)
SYSTEM:	β-globin (A)	yeast/wheat (B-1)	HeLa/Xenopus (B-2)	Tetrahymena (C)
EXCISION SPLICING REACTIONS:	not known	cleavage-ligation: enzymes + ATP required	cleavage-ligation: enzymes + ATP required	2 trans-esterifications: no enzymes, no ATP required*
		phosphate derived from ATP	phosphate derived from 5'end of intron	phosphate derived from 3'end of intron
EXCISED INTRON PROCESSING:	excised 143-nucleotide "lariat" IVS → 130-nucleotide "lariat"; → 143-nucleotide debranched	not known, probably lariat and branched		excised linear IVS → circular IVS + "15-mer" hairpin

*mitochondria require enzymes

APPENDIX

MECHANISMS OF RNA SPLICING AND PROCESSING



3. "Transcription Complexes Make Complex Systems Less Complex"*
ROl GM 22395 (Brown, D.), Carnegie Institute of Washington

One of the excellent systems to study questions of complex regulation is the developing organism. Development is, in essence, gene control. Development requires that the progeny of a single cell experience a spatial pattern and temporal sequence of gene expression and repression, of having genes turned on and off, until the organism contains and maintains its many correct cell types.

Dr. Donald Brown has exploited the development of the frog, Xenopus laevis, to study how the somatic and oocyte cells control gene expression. He has found that the 5S ribosomal RNA gene system in the frog is particularly tractable for research on gene control. The 5S ribosomal RNA genes encode the smallest of the three major RNAs needed to make ribosomes. All haploid Xenopus cells contain 20,000 5S RNA genes, which are composed of two families. The larger family, 98 percent of the total number, are transcribed only in the egg cells and are thus referred to as "oocyte-type" 5S RNA genes. The smaller family contains about 400 genes, 2 percent of the total number. Both oocytes and somatic cells transcribe the genes of the smaller family. For this discussion these will be termed "somatic-type" 5S RNA genes. The oocyte must make and store large amounts of RNA as ribosomes to be used later during development of the embryo. Dr. Brown has chosen to focus on the question: "How is transcription of the oocyte-type genes turned on in oocytes and off in somatic cells?"

In the last five years, investigators have found that gene activity is controlled by specific sequences in DNA which regulate gene transcription. Dr. Brown showed that the regulatory region, or promoter, for 5S RNA gene transcription lies in the middle of each gene, within the coding sequence. The comparable regulatory regions for most genes are located just before the start of the genes.

Meanwhile, Dr. Ronald Roeder found that transcription of the somatic type 5S RNA gene requires at least three protein factors in addition to the transcribing enzyme, which is called RNA polymerase III (pol III). One of these, designated TFIIIA ("transcription factor IIIA ") has a molecular weight of 40,000 daltons. Dr. Brown found that this protein binds to the same sequence that he had identified as the control region for the 5S RNA genes. The other two factors interact with the 40,000 dalton factor and the 5S RNA gene to form a "transcription complex". Pol III, which is not part of the complex, then recognizes this complex and transcribes the gene. The oocyte-type 5S RNA genes require the same three transcription factors and pol III as the somatic-type 5S RNA genes.

One of the characteristics of the transcription complex is its great stability. Many rounds of RNA can be transcribed by a gene without the complex being dissociated. The great stability is the result of the cooperative binding of the several molecules in the complex. The affinity of the components, as a group, to the gene exceeds the affinity of any individual component. The state of the gene is stable not only for many rounds of transcription but through cell division (although during replication orderly perturbations occur).

As part of his experiments to determine what accounts for the turning off, or repression, of the oocyte-type 5S RNA genes in somatic cells, Dr. Brown examined the chromatin of these cells. Chromatin is the naturally-occurring complex of

DNA and its associated proteins. If the DNA is complexed as chromatin in the system, the addition of pol III permits the transcription of the somatic-type 5S RNA genes. However, the oocyte-type 5S RNA genes were not transcribed (even when pol III plus all three transcription factors were added) from their chromatin form.

One of Dr. Brown's group discovered that a chromatin protein, histone H1, blocked the transcription of the oocyte-type genes in somatic cells. When H1 (which is a more mobile histone than the histones H2A, H2B, H3, and H4) is removed, the oocyte-type 5S RNA genes became accessible to transcription. Genes already active, existing as transcription complexes, were resistant to repression by H1. Repression by H1 is a general mechanism which affects several genes while activation requires gene-specific factors.

Somatic-type and oocyte-type genes exist in different states in somatic cells. Both states are stable, but they are composed and maintained by different molecules: the active transcription complex is stable because of the cooperative binding of the several factors to the promoter of the gene; the repressed gene is stable because the histone H1 keeps the transcription factors from getting to the cognate control region on the gene chromatin.

A system in which genes become stably activated or repressed by binding the appropriate proteins "makes the daily life simpler" for the transcribing enzymes of higher organisms. In addition to pol III, higher organisms contain two other polymerases, pol I and pol II. Pol I and pol II also appear to recognize the same transcription complexes as pol III. In contrast, prokaryotes generally recognize specific DNA promoter sequences directly, without utilizing transcription complexes. Since eukaryotic cells have approximately 1000 times as much DNA as prokaryotic cells, eukaryotic polymerases would have a major logistic problem if they had to pick out the right genes to transcribe only on the basis of short DNA sequences. Having the correct genes bound to transcription factors "marks" the relevant genes. Similar, but unwanted, genes are locked up in nucleosomes by the histone.

These results may also help explain how a mature, differentiated cell can maintain its characteristic set of active and repressed genes for long periods of time. Stable complexes are subject to some perturbation during the housekeeping activities of the cell. However, the cooperative binding permits a certain amount of turnover to replace damaged complexes as long as there are a few of the correct replacement molecules available close to the control region (it takes lower concentrations to maintain the complexes than to form them.)

Dr. Brown is now interested in addressing the problem of cell determination, the process by which an embryonic cell becomes "committed" to a particular fate. According to Dr. Brown, what embryologists call "determination" of a gene involves its progression into an active and stable complex. "Commitment" requires the activation of specific genes and the inactivation of others. By use of a transcription complex system, cells can rely on the relative concentrations of various transcription factors to determine which genes will be activated and which will be inactivated. Dr. Brown has found that the concentration of the 40,000 dalton transcription factor of 5S ribosomal RNA genes does correlate with the developmental behavior of these genes. The transcription factor

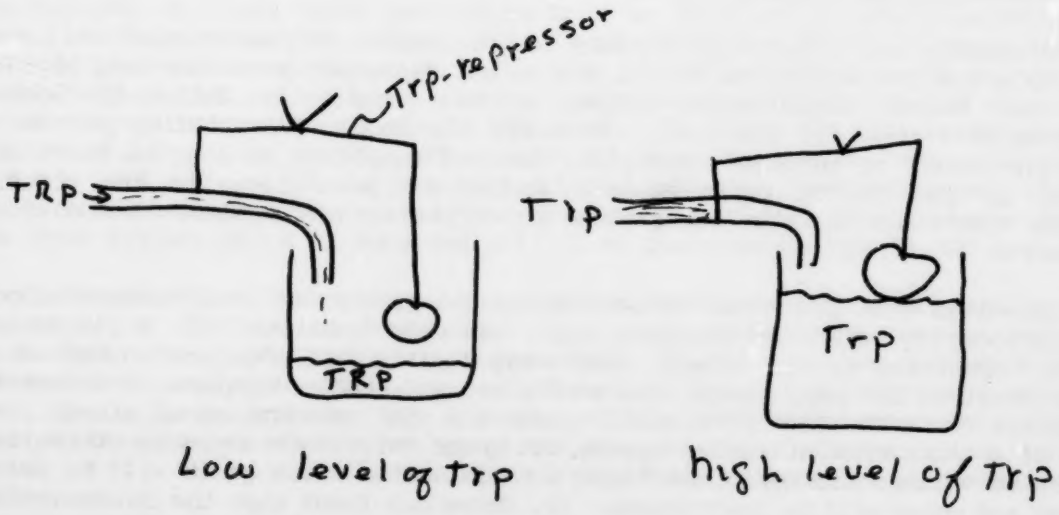
concentration is very high in the oocyte. In the somatic cells of the developing embryo, which have divided several times, the transcription factor has become diluted. In these cells, the somatic-type genes out-compete the oocyte-type genes for the factor and the oocyte-type genes become repressed by binding to H1.

Commitment of a gene by formation of a transcription complex does not necessarily mean it will be expressed. Embryonic cells are committed long before they begin to show their characteristic products. The term "differentiation" refers to the expression of a gene that has already been determined, or activated, but whose transcription is modulated. The determined gene is a stable, but incomplete complex. Other factors will modulate the expression of the gene. The polymerases necessary to read the gene must become available, as well as additional factors, such as hormones. In contrast to higher organisms, bacterial cells differentiate in processes such as sporulation by another mechanism whereby different polymerases, or portions of polymerases, interact with the specific genes.

In Dr. Brown's view, understanding developmental control will result from recognizing basic biophysical principles as they are expressed in cells such as the role of different concentrations of activator versus repressor molecules.

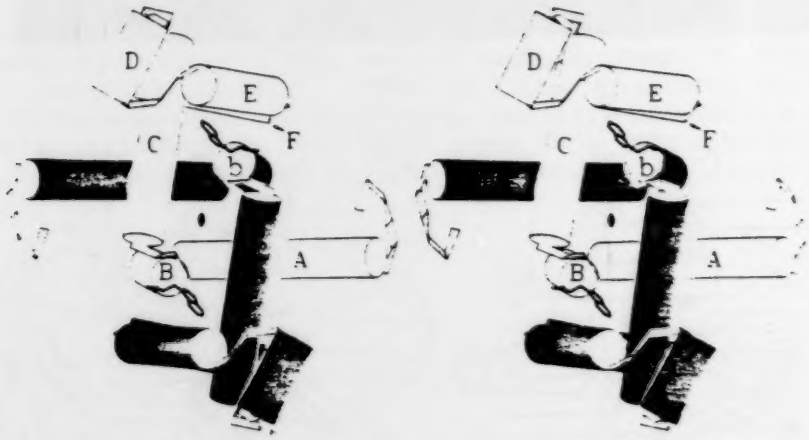
4. "Control of Gene Expression at Atomic Resolution"
R01 GM 15225 (Sigler, P.), University of Chicago

One way enteric bacteria control the amount of the amino acid tryptophan (trp) synthesized in vivo is by regulating the concentration of the enzymes needed for its synthesis through an elaborate system of repression. In the presence of



high concentrations of trp, this amino acid acts as a ligand and binds to a regulator protein, trp-repressor. This amino acid protein complex then binds to an operator site on the bacterial genome which is upstream from the genes coding for the enzymes in the trp biosynthetic pathway. This binding prevents the binding of RNA polymerase, and hence the synthesis of the mRNA necessary for the production of the enzymes. At low concentrations of trp, no binding of trp to trp-repressor occurs, and the affinity of "apo-trp-repressor" for the operator site is sufficiently low that it does not compete effectly with RNA polymerase. Under these conditions, transcription proceeds normally. Thus, trp-repressor acts much like a float valve which allows the amount of trp in a reservoir to control how much more trp flows into that reservoir. That is, as shown above, when the level of trp builds to a certain level, its flow is shut off.

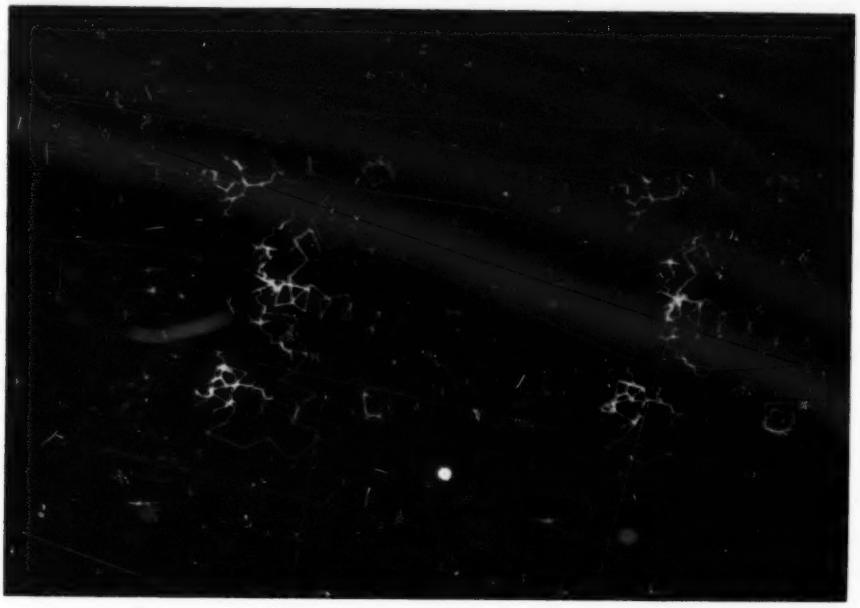
The molecular details of the action of trp-repressor have been the subject of a long time NIGMS grantee, Dr. Charles Yanofsky at Stanford University (GM 09738). These studies represent a classic effort aimed at understanding how a metabolite regulates its own synthesis. Now another NIGMS grantee, Dr. Paul Sigler at the University of Chicago, has provided the first glimpse of how this mechanism works at atomic resolution by determining the crystal structure of the complex between trp and trp-repressor by x-ray crystallography. Dr. Sigler has found that trp-repressor is a dimer with a very unusual subunit structure as shown below.

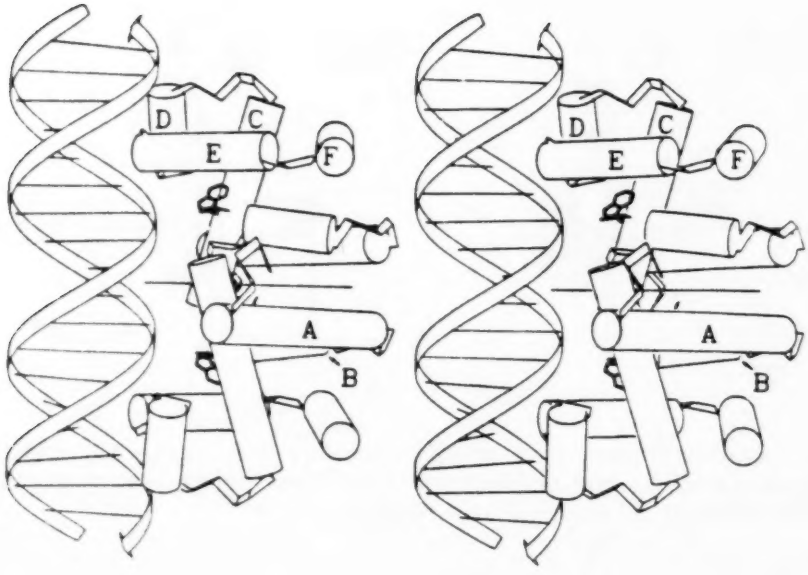


Unlike most multisubunit proteins, which resemble two balls of modeling clay joined by surface contact, the quaternary structure of trp-repressor is formed through the interlocking of some of the helices.

To take this project one step further and try to understand how the trp-repressor binds to operator DNA, Dr. Sigler fitted the protein structure to a structure of B-DNA using a molecular graphics program called FRODO. The structure of the proposed protein nucleic acid complex is shown below. Several features stand out:

1. Both subunits of trp-repressor are involved in the binding to the operator site on DNA. This mode of binding leads to enhanced affinity.



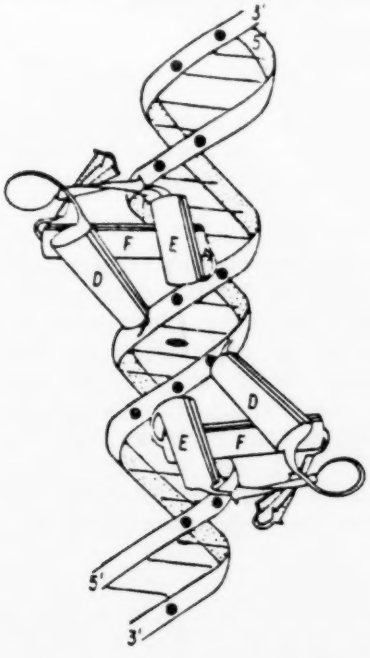


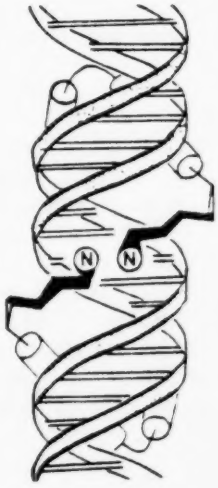
2. The parts of the protein involved in the binding are the D and E helices. These helices are linked by a tight turn, the so-called helix-turn-helix motif common to other DNA binding proteins.

3. The trp is not directly involved in the binding of trp-repressor to DNA but occupies a critical hinge region between the E helix and the remainder of the molecule. It is not difficult to envision that significant changes in the orientation of this critical E helix might take place in the absence of trp.

The completion of the trp-repressor structure brings to a total of 5 the number of DNA binding proteins whose structures have been determined. Four, trp-repressor, catabolite activator protein, CAP, (determined by Dr. Thomas Steitz, Yale University, with support from GM 22778), lambda repressor (determined by Drs. Carl Pabo and Mark Ptashne, Harvard University, with support from GM 29109), and cro repressor (determined by Dr. Brian Matthews, University of Oregon, with support from GM 20066), are regulatory proteins which share certain features discussed below. Of these, only trp-repressor shows regulation by a metabolite, and, hence, is of special significance. The fifth protein, EcoR1, (determined by Dr. John Rosenberg, University of Pittsburgh, with support from GM 25671) is a restriction enzyme whose structure is unique because it is the only example of a DNA binding protein structure determined with a fragment of DNA actually bound to it. Refinement of the data has revealed the structural features which lead to an understanding of the specificity of binding.

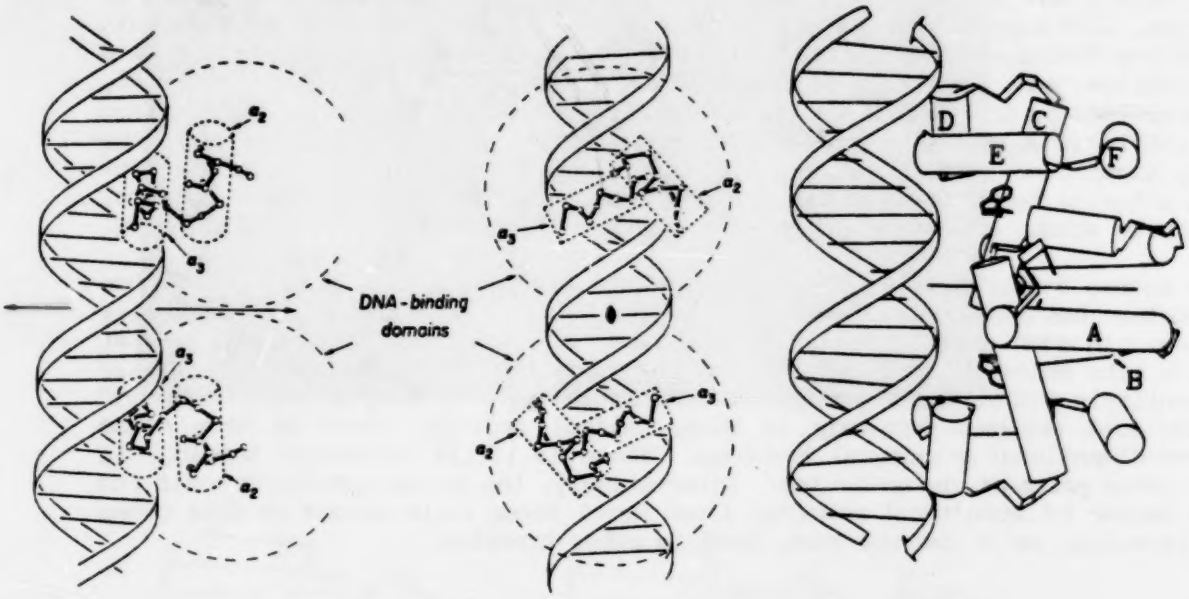
The common feature, which has been found for all regulatory DNA binding, whose structure has been determined by x-ray crystallography, is shown diagramatically below. In each case there are two helices joined by a tight turn. The first helix sits above the DNA near the backbone while the other sits either partially or entirely within the major groove. Although there is only a small amount of amino acid sequence homology in these helical regions, there is very close three-dimensional structural homology. There is little structural homology in the other parts of the molecules. Interestingly, the helix-turn-helix motif (to the degree of structural homology found here) seems to be unique to this group of proteins, as it has not been found in other proteins.





The proposed interaction between the C-terminal domain of CAP and the CAP binding in the lac operon. The helix-turn-helix motif is represented by the E and F helices.

The proposed interaction between lambda repressor's N-terminal arm and the back of the operator site.



Two views of the helix-turn helix motif in the proposed Cro DNA complex.

The proposed interaction between trp-repressor and DNA. The helix-turn-helix is represented by the D and E helices.

The work of Dr. Sigler has provided a first glimpse of how metabolic regulation might occur at atomic resolution. The next step will be to determine the structure of trp-repressor in the absence of trp to find out exactly what changes are responsible for the large change in affinity of repressor for DNA when trp is absent. This aspect is progressing rapidly since Dr. Sigler has crystals which diffract well. In parallel studies, Dr. Oleg Jardetzky at Stanford University (supported by GM 33385) is studying the same system in solution by high resolution NMR. The final chapter will be written when Dr. Sigler is able to confirm the hypothetical structure for the binding of trp-repressor to DNA by solving the structure of the DNA trp-repressor complex.

Genetics Program
Program Organization and Professional Staff - FY 1985

Fred H. Bergmann, Ph.D.	Director
Judith Greenberg, Ph.D.	Deputy Director

Molecular and Medical Genetics Section:

Fred H. Bergmann, Ph.D.	Chief
James Cassatt, Ph.D.	Progam Administrator
Irene Eckstrand, Ph.D.	Program Administrator
Judith Greenberg, Ph.D.	Program Administrator

Genetics of Growth and Differentiation Section:

Judith Greenberg, Ph.D.	Chief
Dorothea Miller, Ph.D.	Program Administrator
Jane Peterson, Ph.D.	Program Administrator
Barbara Williams, Ph.D.	Program Administrator

Support Staff

Maureen Smith	Secretary to the Director
Lucy Clarke	Grants Clerk
Mollie Hilty	Grants Clerk
Bernice James	Grants Clerk
Martha Shanahan	Grants Clerk
Jacqueline Whelchel	Grants Clerk

PHARMACOLOGICAL SCIENCES PROGRAM

OBJECTIVES

The Pharmacological Sciences (PS) Program supports a broad spectrum of research aimed at providing an improved understanding of the biological phenomena and related chemical and molecular processes involved in the actions of therapeutic drugs and anesthetics. The scope of the program ranges from synthetic chemistry and biochemical studies in molecular pharmacology to controlled studies in man. Much of the research supported by the Program integrates such fields as genetics, endocrinology, chemistry, and computer science with the more traditional areas of pharmacological investigation in an effort to characterize molecular interactions and to understand their expression in man.

ORGANIZATION AND STAFFING

Several staffing changes occurred in the PS Program this year. In December, Dr. Christine Carrico was named Director of the Program. At the same time, Dr. M. Janet Newburgh joined the PS Program as a health scientist administrator with responsibility for grants and fellowships in both the bio-related chemistry and basic pharmacology portions of the Program. In June, Dr. James Gilliam retired after 26 years in the Public Health Service.

Currently, Dr. Newburgh and Dr. Carrico handle grants and fellowships in the basic and clinical pharmacology areas. Dr. Newburgh also works with Dr. Carl Kuether to administer the bio-related chemistry section. Dr. Elizabeth O'Hern continues to have responsibility for the anesthesia portion of the Program. All program administrators handle training grants in their respective areas.

RESEARCH OVERVIEW

Research supported by the PS Program is divided into three major areas: anesthesiology, chemistry, and pharmacology. In 1985, approximately $5 million was expended for research in anesthesiology, over $36 million for research in chemistry, and about $20 million for research in pharmacology.

A major portion of the PS Program's budget is expended on research in basic chemistry, in recognition of the importance of such effort to the design, synthesis, and structural elucidation of a wide variety of compounds of potential medicinal and/or practical use. Research in chemistry is multifaceted. A few of the key areas include the design and analysis of model compounds that would shed light on the structure and function of an enzyme's catalytic site; the synthesis of specifically designed artificial enzymes; the synthesis and characterization of specific enzyme inhibitors; ligands that alter nucleic acid function at very specific locations of the genome; and the design and application of new procedures for the efficient synthesis of complex structures, such as antibiotics and anti-tumor agents. One very exciting area of chemical research is the synthesis of nucleotide oligomers that are complementary to key sequences in a viral genome but that also contain a moiety which is capable of reacting covalently with the viral genome to inactivate it. Faster, more efficient

synthetic procedures that are being developed for complex molecules are likely to revolutionize the production of a variety of antibiotics and are making it possible to design related structures that have the desired therapeutic effect but that lack shortcomings characteristic of their natural relatives, such as toxicity or difficulty in delivering the drug to the site of action.

Many plants defend themselves from predators and infective or competing agents by producing toxins. From human society's point of view, such toxins can present problems to individuals or livestock who come into contact, or who eat plant materials containing such toxins. The toxins also are of potential benefit to society if they can be used, or modified for use, as pesticides, antifungals, antibacterials, or anti-tumor agents. An important area of research supported by the PS Program is the biosynthesis and characterization of toxins, as well as identification of the factors or conditions that induce their synthesis.

Active areas of research in pharmacology include: the molecular basis of drug action; drug delivery, absorption, distribution, and elimination; drug metabolism; and factors or substances that modify drug action and drug toxicity. Research activity in both basic and clinical pharmacology is supported. Pharmacological research has progressed to the point that much of the progress during the past year has involved studies at the molecular level. The cytochrome P-450 enzyme family is displaying a diversity that is somewhat reminiscent of immunoglobulin families. The mechanisms by which the synthesis of individual forms of cytochrome P-450 is induced, as well as how the multiple forms of this enzyme family arise, are yielding fascinating information via our current knowledge base and the genetic, biochemical, and molecular biological tools currently available to us. These enzymes are interesting not only because of their effects on xenobiotics, but also because they appear to be involved in normal metabolism of cholesterol, fatty acids, prostaglandins, leukotrienes, bile acids, and sterol hormones.

Another focus of considerable pharmacological research effort is hormone-sensitive adenylate cyclase, a system that directly or indirectly is affected by several drugs. This system has turned out to be considerably more complex than previously thought but is yielding to solution and to understanding. One fascinating finding is the similarity among guanine nucleotide-binding regulatory proteins that serve as information transducers for these receptors, as well as their similarity to at least one oncogene product.

Current NIGMS-supported research in anesthesiology is directed toward obtaining the basic knowledge for the development and improved use of safer anesthetic agents. It includes a major effort directed toward a better understanding of the mechanisms of anesthetic action at the molecular, endocrine, and cellular levels. Other major areas of interest are studies of toxic or side effects of anesthesia, the metabolism of anesthetics, and how effects or the metabolism of anesthesia may be altered in specific cases, for example in burned patients or during surgery. Anesthesia may significantly affect a patient's reaction to treatment. For example, certain common forms of anesthesia alter differentially the intracranial pressure changes that occur in patients with head trauma who are being treated by hyperventilation. Information regarding these effects clearly is important in management of these patients. Other research is directed at understanding toxic side effects of anesthesia and identifying those

individuals who have a greater than normal tendency for adverse reaction to anesthesia. For example, one project is seeking a better understanding of the molecular mechanism of malignant hyperthermia and how to determine whether a given individual is susceptible to this condition.

In April, the PS Program sponsored a symposium on "Novel Genetic Approaches in Molecular Pharmacology" as part of the spring meeting of the American Society for Pharmacology and Experimental Therapeutics in Anaheim. The thought-provoking presentations covered a range of contemporary research topics in pharmacological sciences at this well-attended symposium.

The 20th anniversary of the PS Program was celebrated in March with a reception to which all Pharmacology Research Associate Program (PRAT) preceptors, as well as present and former PRAT fellows, were invited. Also part of the celebration was a lecture by Dr. Alfred Gilman, a former PRAT fellow, and a poster presentation of the exciting and varied research projects being undertaken by the current PRAT fellows at NIH.

RESEARCH TRAINING

Pharmacological scientists serve important roles in all areas of scientific and health-related endeavors in academia, government and industry. To provide a continuing source of such scientists is the purpose of the training programs of the PS Program. Since 1975, research training in the pharmacological sciences and clinical pharmacology has been supported under the authority of the National Research Service Act. As of 1983, research training in anesthesiology has also come under the auspices of the PS Program. Multidisciplinary predoctoral training programs in the pharmacological sciences are supported by institutional fellowships, as is postdoctoral research training in anesthesiology and clinical pharmacology. Individual fellowships also are available for postdoctoral training in all of the above fields.

The multidisciplinary nature of the institutional predoctoral training programs in the pharmacological sciences provides the student with a combination of skills and techniques necessary to address a wide spectrum of pharmacological problems. Programs are designed to offer students the opportunity to apply knowledge of molecular biology, medicinal chemistry, neurobiology, biochemistry, physiology, and other related biological sciences to problems of pharmacological interest. These predoctoral programs also are available to students in disciplines other than pharmacology. Regardless of their departmental affiliation, all trainees receive a thorough foundation in pharmacological principles and their application to all aspects of biomedical research. In FY 1985, there were 38 institutional predoctoral training programs, supporting 285 trainees.

The training of M.D.'s in research remains a strong commitment of the National Institute of General Medical Sciences (NIGMS) and the PS Program. Support is provided through various funding mechanisms. Institutional postdoctoral fellowship programs in clinical pharmacology and anesthesiology are designed to train physicians in the application of laboratory research principles to the study of these disciplines. In FY 1985, there were 9 clinical pharmacology training programs supporting 28 trainees and 5 anesthesiology training programs supporting 8 trainees. Individual postdoctoral fellowships also are available in

these disciplines. Research career development award applications are accepted by the Program only from those investigators who have expressed a commitment to research careers in clinical pharmacology or anesthesiology. The PS Program currently supports two such investigators in anesthesiology and one in clinical pharmacology.

The PRAT Program is the only intramural activity of NIGMS. This unique Program provides opportunities for 22 highly qualified postdoctoral fellows to spend 2 years in the intramural laboratories of the National Institutes of Health and the Alcohol, Drug Abuse, and Mental Health Administration. There are currently 86 formally appointed preceptors whose interests include pharmacokinetics, hormone-receptor interactions, cyclic nucleotide biochemistry, molecular toxicology, and virtually all aspects of the neurosciences. In addition to the extramural support for clinical pharmacology, the PRAT Program provides an individualized curriculum for interested and qualified physicians to pursue research training in clinical pharmacology at the National Institutes of Health.

RESEARCH HIGHLIGHTS

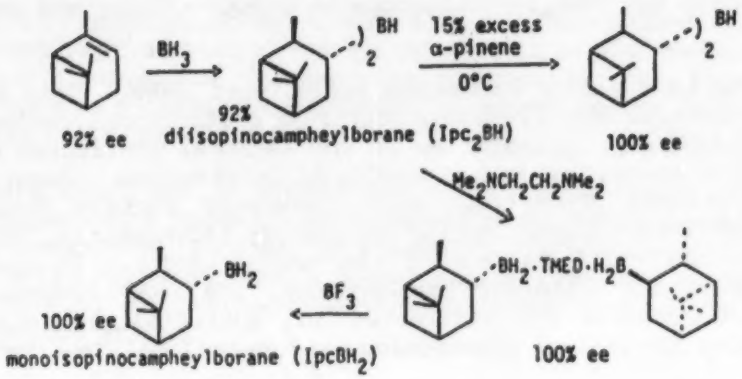
"Asymmetric Synthesis via Chiral Hydroboration"
R01 GM 10937-22 (Brown, H.), Purdue University

Dr. Herbert Brown and his coworkers have recently discovered a general method of asymmetric synthesis for preparation, in optically pure form, of almost any chiral compound. The method involves hydroboration, which Dr. Brown discovered and developed and for which he was awarded the Nobel prize in chemistry for 1979.

Two asymmetric hydroborating agents are very commonly utilized: diisopinocampheylborane (Ipc_2BH) and monoisopinocampheylborane ($IpcBH_2$). Both reagents are readily obtained from α-pinene, a cheap and readily available material prepared from oil of turpentine. Commercial α-pinene is only 92 percent optically pure. However, if the crystalline Ipc_2BH formed upon hydroboration of α-pinene (92 percent enantiomeric excess) is allowed to stand at 0°C in the presence of a slight excess of α-pinene, the major isomer becomes incorporated into the crystalline reagent, leaving the unwanted minor isomer in solution.

Because the reaction commonly proceeds rapidly to the dialkylborane stage, the monoalkylborane must be prepared by an indirect route. Treatment of diisopinocampheylborane with one-half molar equivalent of N,N,N',N'-tetramethylethylenediamine (TMED) provides the 2:1 adduct, $TMED\text{-}2BH_2Ipc$, which crystallizes out in enantiomerically pure form. Pure monoisopinocampheylborane then is readily liberated by treatment with boron trifluoride. (See Figure 1)

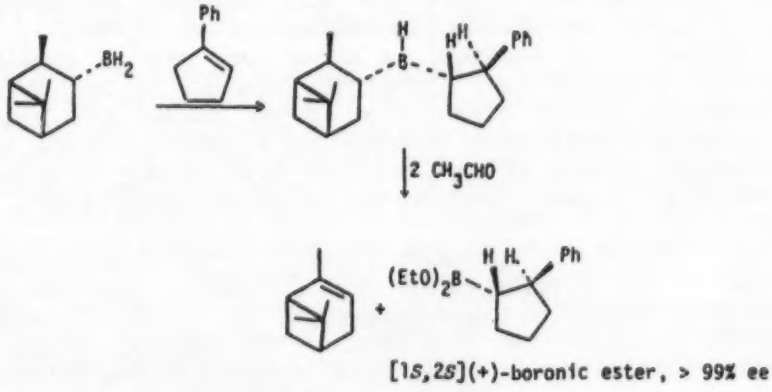
Figure 1



Preparation of Ipc_2BH and $IpcBH_2$

Three recent developments in Dr. Brown's laboratory are of major importance. First, is the discovery that treatment of the asymmetric hydroboration products with acetaldehyde removes the isopinocampheyl group as α-pinene, yielding the optically active boronic esters.

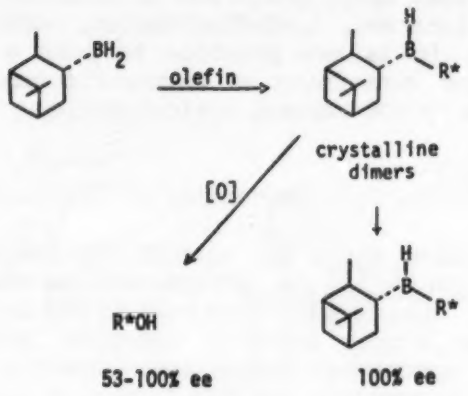
Figure 2



Synthesis of Optically Pure Boronic Esters

Second, the reaction products from monoisopinocampheylborane and hindered prochiral alkenes, IpcR*BH, exist as crystalline dimers. Crystallization provides a product of essentially 100 percent optical purity.

Figure 3

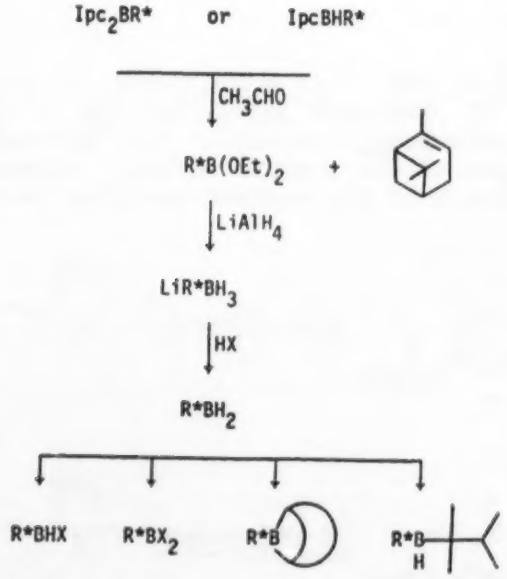


Preparation of IpcR*BH of Essentially 100% ee

This makes it possible to prepare many chiral boronic esters of essentially 100 percent enantiomeric excess.

Third, it is readily possible to convert these boronic esters into the corresponding monoalkylborohydrides and, by treatment with acid, obtain the optically active monoalkylborane. The optically active borane can be converted by known methods into all of the organoborane reagents which have been found to be valuable in syntheses.

The whole process is summarized below.



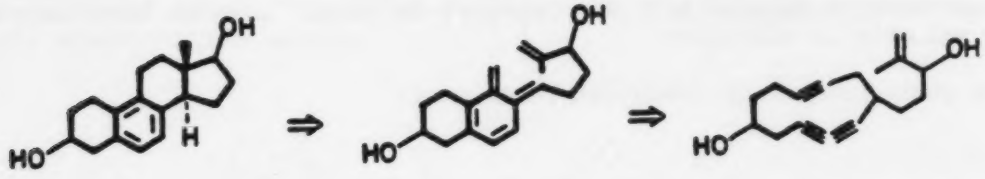
Synthesis of Boron Intermediates of 100% ee

It is well established that alkyl groups can be transferred from boron to essentially any element of interest, including carbon, with complete retention of stereochemistry. Thus, it is now possible to plan a rational synthesis of practically any compound containing an asymmetric center, either the (+) or (-) isomer, in essentially 100 percent optical purity.

"Organometallic Chemistry"
R01 GM 22479-10 (Vollhardt, K.), University of California, Berkeley

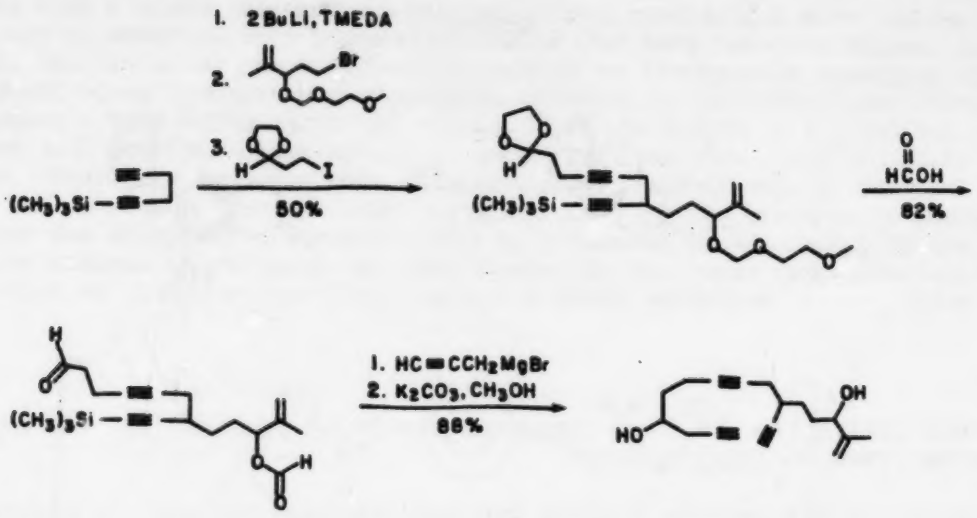
Dr. Vollhardt has developed three new methods for the total synthesis of steroids that use transition metals as catalysts or as reagents. The first procedure uses catalytic dicarbonylcyclopentadienylcobalt [$CpCo(CO)_2$] to fuse the ABC-portion of the A-ring aromatic steroids onto an already existing D-ring. The other two approaches employ stoichiometric amounts of dicarbonylcyclopentadienylcobalt to annelate the BCD-portion of the steroids to an already existing A-aromatic ring. Significant chemo-, regio-, and stereoselectivity is achieved in these transformations. Dr. Vollhardt and his coworkers have now found an even more direct approach to this physiologically important class of molecules, in which an enetriyne is cyclized with catalytic cobalt in one step to assemble all four rings. The retrosynthetic analysis is shown in Scheme I.

Scheme I



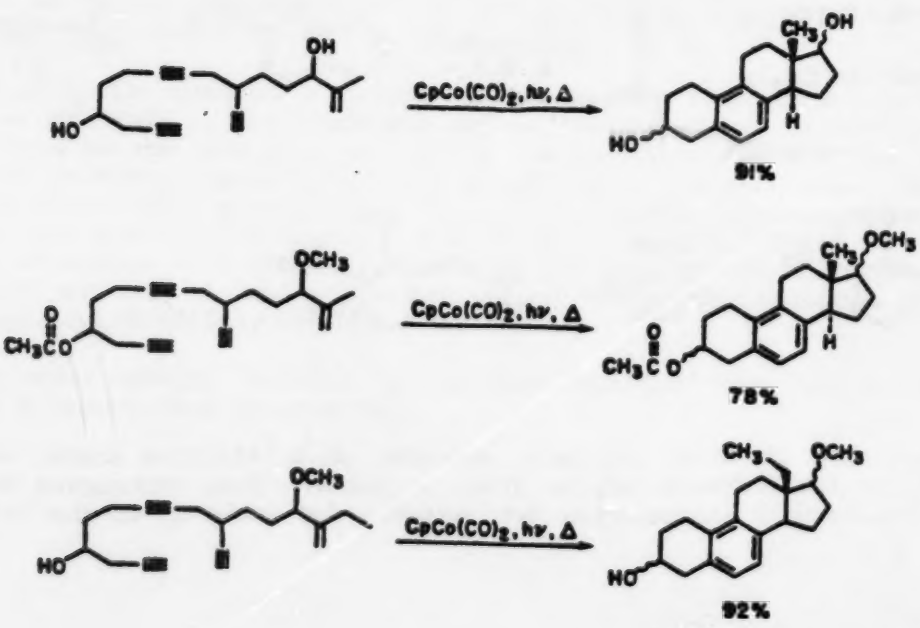
This cyclization gives rise to an extraordinarily rapid assembly of B-aromatic steroids, a class of compounds which was not previously accessible. The approach should be extremely versatile for synthesizing analogs and novel medicinal agents based on the steroid nucleus. Scheme II depicts an efficient and convergent synthesis of the required enetriyne percursor.

Scheme II



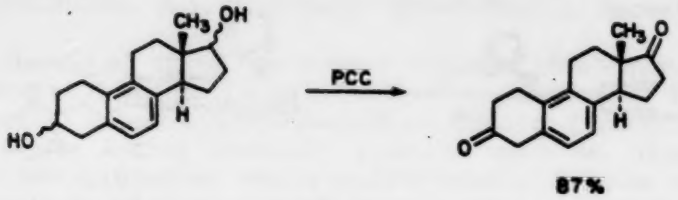
The examples in Scheme III demonstrate the efficiency and versatility of the method. Target molecules are assembled with complete stereoselectivity with respect to the CD-ring junction, which emerges trans, as in the naturally occurring systems.

Scheme III



The identity of the cyclization products was ascertained by spectral and chemical means. Thus, as shown in Scheme IV, oxidation of the B-ring aromatic 3,17-diol gives the corresponding known diketone.

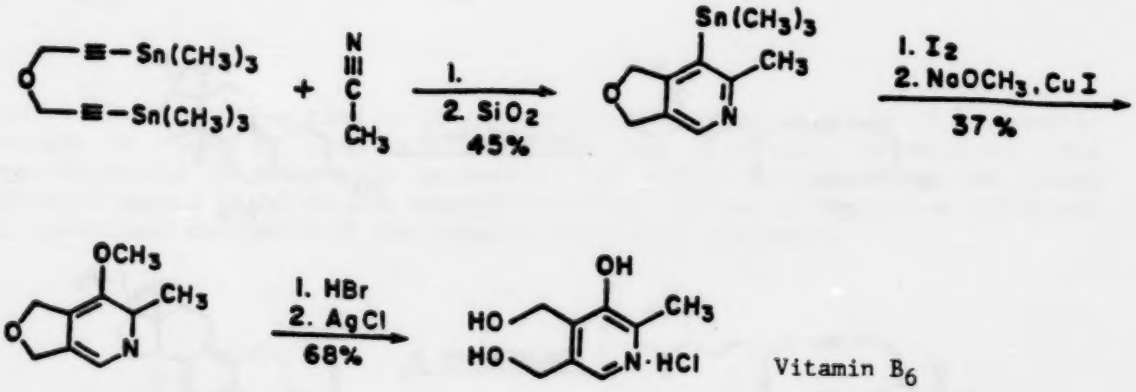
Scheme IV



Several derivatives of these compounds are being prepared for testing in an androgen receptor binding assay.

By utilizing his previous finding that cyclizations catalyzed by dicarbonylcyclopentadienylcobalt are not restricted to unsaturated moieties containing only carbon, Dr. Vollhardt has developed a highly chemo- and regioselective construction of vitamin B_6 (Scheme V).

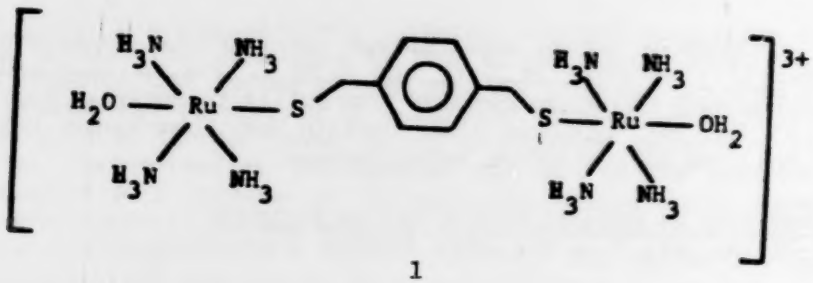
Scheme V



A conceptually identical approach, executed on a kilogram scale, was used recently by the Hoffmann-LaRoche group in Basle. Thus, strategies developed in Dr. Vollhardt's laboratories are suitable for scale-up to the industrial level.

"Through-Space Electron Transfer Reactions"
R01 GM 31383-03 (Lewis, N.), University of New Brunswick

For more than a decade, researchers have speculated that sulfur atoms and benzene rings play an important role in metalloproteins that have redox functions. Based on this, Dr. Lewis has prepared models designed to incorporate molecular fragments found along the presumed electronic pathways in proteins. For example, in compound 1, the sulfur atoms and benzene rings are placed in a rigid environment and are situated close enough to each other so that their orbitals can overlap. There are two ways that an electron can pass through a molecule like this. The first is a "through-bond" transfer, in which the electron is presumed to follow the molecular framework. This is prevented in compound 1 by the two methylene groups, which serve as insulators. It has never been conclusively proven that an electron can pass through a fully saturated carbon like the

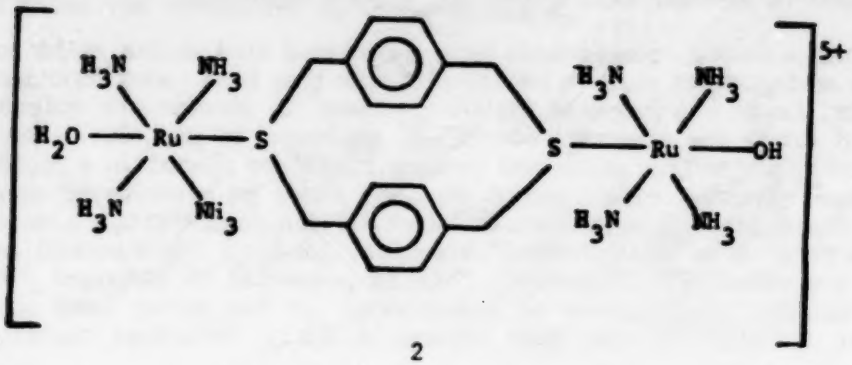


1

two methylenes, in an unstrained environment, at any measurable rate. Thus, the electron probably is forced to use the only other possible mechanism available to it, which is the "through-space" mechanism. "Through-space" transfer is so named because the orbitals used in transferring the electron belong to atoms that are not connected directly to each other. This is thought to be the type of transfer that occurs in redox proteins.

In compound 1, the electron is thought to pass directly from a lone pair orbital on the sulfur atom to the p-orbitals of the benzene ring. It then probably passes along to the lone pair orbital on the second sulfur atom and finally leaves the bridging ligand by entering an orbital on the metal atom. However, the intervalence band corresponding to a through-space transfer is ten times narrower than predicted by theory. This is tentatively attributed to coupling of this transition to a vibrational mode of the ligand. The electrochemistry (E_f = 0.27 and 0.37 V. vs. a standard calomel electrode) indicates that the metal centers are strongly coupled electronically.

A second model system, compound 2, has been developed that has two benzene rings in a face-to-face orientation.



2

This molecule apparently shows simultaneous two-electron transfers on the cyclic voltammetric time-scale at 0.11 and 0.28 V. (vs. a standard calomel electrode), taking each ruthenium formally from Ru(II) to Ru(IV). Such behavior is unprecedented and may indicate that certain metalloproteins having this feature could function as 2-electron redox systems.

"Bioorganic Chemistry of Nucleic Acid Alkylphosphonates"
R01 GM 31927-02 (Miller, P.), The Johns Hopkins University

Dr. Paul Miller of The Johns Hopkins University is seeking to develop oligonucleotide analogs that can control gene expression at the level of DNA transcription or at the level of mRNA translation or processing. Toward this goal, he has used solid phase techniques to synthesize oligodeoxyribonucleoside methylphosphonates of defined base sequence that are specifically designed to be taken up intact by mammalian cells in culture and to form stable, hydrogen-bonded complexes with complementary target DNA or RNA sequences.

Oligomers up to 15 nucleoside units long have been synthesized and characterized by procedures developed by Dr. Miller. He has found that oligomers complementary to the splice junction sequences of SV40 precursor mRNA selectively inhibit virus protein synthesis in infected cells. Herpes simplex virus type 1 DNA synthesis, protein synthesis, and virus growth are selectively inhibited in virus-infected Vero cells by an oligomer which is complementary to the exon/intron junction of an immediate early mRNA.

Dr. Miller recently designed a new type of methylphosphonate oligomer that forms covalent crosslinks with its target nucleic acid. The oligomer contains a derivative of psoralen attached to its 5'terminal nucleotide residue which, when irradiated with 365 nm light, forms adducts with cytosine, uridine or thymine residues of the complementary nucleic acid strand. In preliminary experiments utilizing very low concentrations of psoralen-derivatized methylphosphonate oligomer complementary to β globin mRNA, the compound formed covalent complexes between the complementary oligomer and mRNA after irradiation, resulting in a 75 percent inhibition of translation of the mRNA. As the specificity of the oligodeoxyribonucleoside methylphosphonates resides in their ability to bind to complementary nucleic acid sequences, it should be possible to use nucleic acid sequence information to design novel derivatives of

methylphosphonate oligomers, which can be used to study gene expression and the function of specific proteins in normal, transformed or virus-infected cells.

"Host Molecules that Complex and Catalyze"
R01 GM 12640-21 (Cram, D.), University of California, Los Angeles

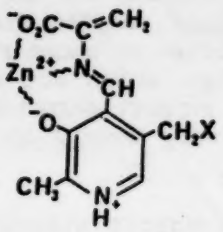
A powerful probe for learning about the physicochemical basis of health is the creation of designed organic chemical models for chemical reactions moderated by enzymes. Transacylases are important in the synthesis and degradation of proteins and fats. The active sites of transacylase enzymes from a variety of organisms appear to contain concave complexing sites on whose rims are gathered an hydroxyl, an imidazole, and a carboxyl group, which act cooperatively to catalyze acyl transfer reactions. The peptide support structures that provide organization to the transacylase active sites vary widely in their molecular weights and amino acid sequences.

Based on the hypothesis that entirely different support structures can be used to synthesize transacylase catalysts as long as the active sites are organized similarly to those of the enzymes themselves, Dr. Cram has designed a potential artificial transacylase consisting of a binding site and the three functional groups of the natural transacylases in an arrangement resembling that of the natural enzyme. He devised an incremental approach to its synthesis to test the performance of each portion of the molecule before fully assembling the four parts. The binding site was successfully constructed and found to collect and orient the reacting molecules as anticipated. The hydroxyl group that accepts the acyl group was next put in place and was found to perform its assigned task of acting as an acyl acceptor. Most recently, in a 30 step synthesis, the imidazole group was properly located in the desired synthetic catalyst with dramatic results. The compound acted as an acyl acceptor in the absence of an external base and the observed acetylation reaction rate was 100,000 times faster than that measured with the various catalytic parts unassembled. Although the final carboxyl group has not yet been introduced into the compound, the acylated catalyst undergoes slow hydrolysis to provide catalytic turnover. Dr. Cram hopes that addition of this carboxyl group will accelerate the hydrolysis.

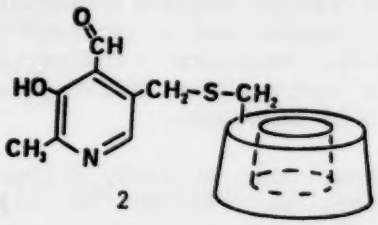
"A Mimic of Tryptophan Synthetase"
R01 GM 18754-01 (Breslow, R.), Columbia University

In tryptophan synthetase, a pyridoxal cofactor binds a serine residue and promotes the formation of a dehydrated derivative that reacts with indole, held in a binding pocket, to form tryptophan. Dr. Breslow and his coworkers in the Department of Chemistry at Columbia University have synthesized a cyclodextrin molecule carrying a pyridoxal unit (structure 2) that imitates the ability of tryptophan synthetase to bind two components and assemble them, forming a new carbon-carbon bond. In the artificial enzyme, the pyridoxal unit binds an amino acid which forms the aminoacrylic acid derivative compound 1, while the cyclodextrin unit functions as a binding pocket to hold the indole in position to react with the amino-acrylic acid derivative. The system will catalyze a reaction starting with serine but is more effective when the starting material is chloroalanine. The latter compound generates the aminoacrylic acid intermediate 1 more readily with fewer side reactions. Evidence for this mechanism is based on the fact that the cyclodextrin-pyridoxal catalyst gives

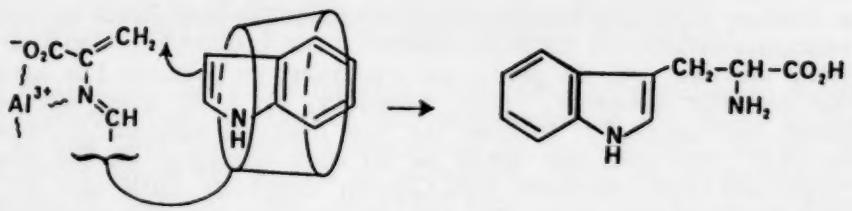
higher yields than does a simple pyridoxal derivative, and also on the fact that the product tryptophan has some optical activity. This suggests that the product is formed inside the chiral cyclodextrin unit. Dr. Breslow is attempting to improve this catalyst by adding an additional basic catalytic group and eliminating some of the flexibility in the system.



1



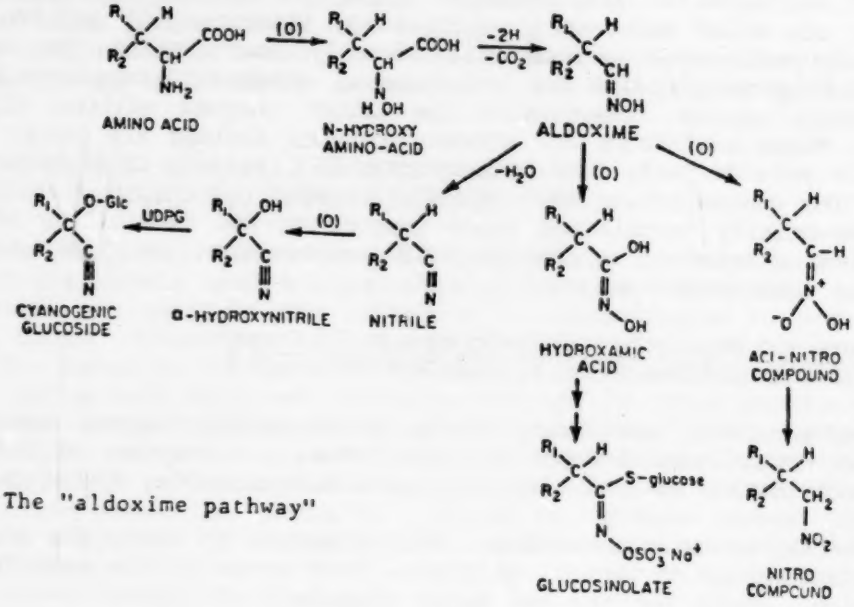
2



"Metabolism of Aromatic Compounds"
R01 GM 05301-28 (Conn, E.), University of California, Davis

In many higher plants, the aromatic amino acids phenylalanine and tyrosine are converted to cyanogenic glycosides by the biosynthetic pathway shown below. Many other higher plants use this same pathway for converting the aliphatic amino acids, valine, isoleucine and leucine, to cyanogenic glycosides. (Some animals, including millipedes and zygaenid moths, also utilize the pathway to produce cyanogens.) While studying the formation of an unknown cyanogenic compound from tyrosine in cell suspension cultures of California poppy (Eschscholtzia californica), Dr. Conn and his colleagues observed the formation of 1-(4'-hydroxyphenyl)-2-nitroethane. This nitrocompound is produced by cells which are osmotically stressed by 0.5 M sorbitol, a condition which is known to stimulate the production of isoquinoline alkaloids by the same cells. Microsomal preparations from stressed cells catalyze the conversion of either L-tyrosine or 4-hydroxyphenylacetaldoxime to 1-(4'-hydroxyphenyl)-2-nitroethane in the presence of NADPH. The aci-nitrocompound, produced by N-hydroxylation of the aldoxime, is tentatively postulated as an intermediate in the formation

of 1-(4'-hydroxyphenyl)-2-nitroethane. Thus, a third class of toxic natural plant product—compounds containing an aliphatic nitro group—shares with cyanogenic glycosides and glucosinolates the initial steps of a biosynthetic pathway involving N-hydroxylation of amino acids. This class includes β-nitropropionic acid, which occurs as a glucose ester in two genera of legumes (Astragalus and Indigophera) and is responsible for the toxicity of several species to livestock.



The "aldoxime pathway"

"Selective Destruction of Cytochrome P-450 by Drugs"
R01 GM 25515-07 (Ortiz de Montellano, P.), University of California, San Francisco

The cytochrome P-450 family includes some forms of the enzyme primarily devoted to the metabolism of foreign compounds and other forms primarily devoted to the processing of endogenous substrates, including cholesterol, fatty acids, prostaglandins, leukotrienes, bile acids, and sterol hormones. Organisms are known to be much more vulnerable to oil-soluble toxins when their P-450 shield is impaired, but little information is available on whether the loss of P-450 itself, in the absence of a toxic challenge, is deleterious. To answer this question, it is necessary to eliminate the P-450 enzymes involved in drug metabolism without eliminating the P-450 enzymes that are an integral part of essential biosynthetic processes, such as sterol biosynthesis. This need is met by 1-aminobenzotriazole, a suicide substrate developed in the laboratory of Dr. Ortiz de Montellano that irreversibly inactivates up to 80 to 90 percent of the hepatic cytochrome P-450 in uninduced rats without detectably affecting any of the sterol biosynthetic enzymes so far tested. The investigator has found no detectable changes in a variety of parameters that measure toxicity or in the apparent health of the animals after administration of the agent to rats for up to two weeks. Thus, drug metabolizing forms of cytochrome P-450 appear to be important primarily because of the protection they provide against

xenobiotics. 1-Aminobenzotriazole should find wide utility in studies of the role of cytochrome P-450 enzymes in physiological and toxicological processes.

In order to examine more precisely the metabolic role of a single form of cytochrome P-450, Dr. Ortiz de Montellano has been attempting to differentially inactivate the closely related enzymes that hydroxylate fatty acids, prostaglandins, and leukotrienes. He has found that incorporation of an acetylenic group into the substrates of these different enzymes results in their highly specific, differential inactivation. Thus, the enzyme that hydroxylates the terminal carbon of short chain fatty acids (lauric acid) can be inactivated by 10-undecynoic acid without significantly inactivating the enzymes that hydroxylate prostaglandins and leukotrienes, whereas 17-octadecynoic acid, the longer chain analog, inactivates the latter enzymes without affecting the former. These inhibitors and closely related analogs are being utilized to study the role of fatty acid hydroxylases in a variety of physiological processes. The demonstration that closely related cytochrome P-450 enzymes can be differentially inactivated lends support to the feasibility of developing inhibitors of specific cytochrome P-450 enzymes that could be used as therapeutic or insecticidal agents.

"Structure and Regulation of Cytochrome P-450 Genes"
R01 GM 30701-04 (Adesnik, M.), New York University

The broad substrate specificity of the cytochrome P-450 system results from the existence of at least a dozen distinct forms, or isozymes, of P-450, many of which individually exhibit broad, albeit distinguishable, specificity profiles.

Dr. Adesnick is using recombinant DNA techniques to study the structures and interrelationships of various cytochrome P-450 genes and the mechanisms of their regulation. cDNAs for the two major phenobarbital-induced P-450 isozymes of rat liver (P-450b and P-450e), the corresponding gene for one of them (P-450e), and five other related cross hybridizing genes have been cloned. Detailed sequence analysis of portions of all these cloned genes provides evidence for extensive sequence exchange within the region containing exons 7 and 8 (there are 9 exons altogether) and the intron between them by a gene conversion mechanism. Based on these data and published sequence data for other P-450 isozymes, Dr. Adesnick has proposed gene conversion to be an important evolutionary mechanism for generating genes encoding related cytochrome P-450 isozymes.

The P-450b and P-450e genes are expressed and induced by phenobarbital in the adrenal gland and liver but not in the kidney, brain or preputial gland. Low levels of expression of the P-450b and/or P-450e genes were found in the lung. These levels were approximately 10 fold less than the basal level in the liver and were unaffected by phenobarbital administration. A related gene that encodes an as yet unidentified P-450 isozyme was expressed in the preputial gland, an organ which actively synthesizes steroids. The expression of several P-450 genes in organs that are active in steroid synthesis, as well as the known capacity for P-450b and P-450e to metabolize steroids, suggest steroids as the endogenous substrates for the enzymes encoded by this gene subfamily.

Immunoscreening of cDNA libraries with antibodies raised against P-450 PB-1 and 2c, which are major constitutive forms in rat liver, has permitted isolation of cDNA clones for three distinct but related cytochromes P-450, which on the basis

of DNA sequence analysis encode proteins that are approximately 50 percent homologous to cytochrome P-450b. Differential hybridization selection experiments indicate that one of the clones identified with the anti-PB-1 antibody appears to correspond to the mRNA encoding the immunogen, whereas the second corresponds to an immunochemically related but heretofore uncharacterized isozyme. Hybridization experiments show that the mRNAs corresponding to one of the clones is induced 8-16 fold by phenobarbital administration, whereas the second is induced only 2-4 fold. Given the fact that these clones represent genes of the same family as the P-450b and P-450e genes and that genomic Southern blotting with these probes gives non-overlapping patterns which are as complex as that obtained with P-450b or P-450e probes, this gene family appears likely to contain 20 to 30 family members that are divided into at least two subfamilies.

"Alteration of Metabolism by Structural Modification"
R01 GM 25418-07 (Nelson, S.), University of Washington

Dr. Nelson has previously shown that N-acetyl-p-benzoquinone imine is a toxic metabolite of the widely used analgesic-antipyretic, acetaminophen, and that cytochrome P-450 can catalyze the oxidation of acetaminophen to N-acetyl-p-benzoquinone imine. Although cytochrome P-450-mediated oxidation is largely responsible for formation of the toxic metabolite in the liver, prostaglandin synthase may be the more important oxidative catalyst in other tissues such as the kidney. Ram seminal vesicle microsomes, a rich source of prostaglandin synthase, catalyze the formation of both N-acetyl-p-benzoquinone imine and radical recombination products, indicating the presence of potentially toxic 2- and 1-electron oxidation products. Inhibition studies showed that the hydroperoxidase component of prostaglandin synthase is responsible for the activation of acetaminophen.

In addition to its activity as a cosubstrate of prostaglandin synthase, acetaminophen stimulates the conversion of prostaglandin G_2 to prostaglandin H_2 at relatively low concentrations (0.2mM), whereas it inhibits the formation of prostaglandins from arachidonic acid at high concentrations (>10mM). A series of analogs of acetaminophen have this effect; the order of potency for both stimulation of prostaglandin H_2 synthesis and inhibition of prostaglandins is correlated with the electrochemical half-wave potentials of the compounds. Therefore, it may be difficult to prepare an analog of acetaminophen that is a good inhibitor of prostaglandin synthase, and thus retains therapeutic activity, but that is less capable of being oxidized to reactive metabolites.

Sometimes it is useful to generate reactive metabolites. For example, aromatase is a cytochrome P-450 enzyme complex which catalyzes the oxidation of the male hormonal androgens to the female hormonal estrogens. Thus, inhibition of aromatase can limit estrogen production which is of therapeutic benefit in treating cancers that are dependent on estrogens for growth. Dr. Nelson has made minor structural modifications to the hormonal steroid, androstenedione, in attempts to develop a suicide inhibitor for the enzyme. His research group has prepared a 10-mercapto analog, which is isosteric with androstenedione, and a 19-mercapto analog, which is isoelectronic with the first oxidation product formed by aromatase oxidation of androstenedione. Both analogs are potent suicide inhibitors of human placental aromatase, probably via their oxidation to reactive sulfenic acids that bind to the active site of the enzyme.

"Interindividual Variability in Oxidative Metabolism in Man"
P01 GM 31304-03 (Wilkinson, G.), Vanderbilt University

Differences in the quantitative handling of a drug by oxidative metabolism often are major contributors to the sometimes wide interpatient variability in drug responsiveness. Patients who metabolize a drug readily may receive ineffective therapy from the administration of usual drug dosages, whereas poor metabolizers may have an increased risk of adverse effects, or when active metabolites are involved, a reduction in efficacy. Dr. Wilkinson and colleagues at Vanderbilt University have continued their studies of factors involved in such variability with particular interest in genetically determined polymorphisms involving the 4-hydroxylation of debrisoquine and mephenytoin. In collaboration with Japanese investigators at the University of Tsukuba, they have compared the frequency of impairment of these two independently inherited pathways. In Caucasians living in middle Tennessee, 8 percent were found to be "poor metabolizers" of debrisoquine, an incidence similar to that found in other populations of the same ethnic origin. In contrast, no individual was found in Japan who had this phenotype in over 100 consecutive, unrelated subjects. On the other hand, the frequency of impaired oxidation of the S-enantiomer of mephenytoin was sixfold higher in Japanese versus Caucasians, that is 18 percent and 3 percent, respectively. These pronounced interethnic differences in oxidative metabolism confirms a genetic independence of the two polymorphisms and supports anecdotal and clinical experience that drug responsiveness often differs substantially between ethnic groups such as Orientals and Caucasians. Given the international nature of drug development and use and the multiracial populations of many countries, the findings indicate the need for increased recognition of such differences in drug evaluation.

While the debrisoquine oxidative status is now known to be linked to the metabolism of over 30 different drugs ranging from tricyclic antidepressants through beta-adrenergic blockers to antiarrhythmic agents, much less is known about the specificity of the S-mephenytoin polymorphism. In collaboration with Dr. F. P. Guengerich, Dr. Wilkinson has identified a number of other anticonvulsant drugs whose metabolism is thought to be controlled by the same genetic factor(s) as that of mephenytoin. This grouping was based on their ability to competitively inhibit the 4-hydroxylation pathway in human liver microsomes *in vitro*. A common feature for potent inhibition is the presence of an α-aryl-N-alkyl lactam substructure, *e.g.*, ethotoin, methsuximide and mephobarbital. Clinical studies with mephobarbital confirmed the *in vitro* prediction that its 4-hydroxylation co-segregates with that of mephenytoin. These invesigators also isolated, purified, and characterized a specific cytochrome P-450 isozyme from human liver microsomes that is primarily responsible for the 4-hydroxylation of mephenytoin. The isozyme differs catalytically, electrophoretically and immunochemically from other human oxidative isozymes. Such isozymes will be useful for studying the molecular and genetic regulation of cytochrome P-450 function in humans.

The importance of a major gene effect in drug metabolism also was demonstrated during the clinical evaluation of a new antiarrhythmic agent, propafenone. Early studies indicated that the drug is efficacious but exhibits pronounced interpatient variability in disposition and effective plasma concentrations. Dr. Wilkinson and colleagues have now shown this to be due to the production of

a pharmacologically active metabolite, 5-hydroxypropafenone, which is determined by the debrisoquine phenotype of the patients. Thus, poor metabolizers have unexpectedly high plasma levels of unchanged drug and reduced concentrations of active metabolite, leading to an increased risk of adverse effects unless the interphenotypic difference in metabolism is taken into account.

"Regulation of Cyclic Nucleotide Metabolism"
R01 GM 34497-05 (Gilman, A.), University of Texas Health Sciences Center, Dallas
"Molecular Basis of Receptor-Cyclase Coupling"
R01 GM 27800-05 (Bourne, H.), University of California, San Francisco

Members of a family of membrane-bound guanine nucleotide-binding regulatory proteins (G proteins) serve as information transducers; they act as coupling factors to transmit information from receptors for a variety of regulatory compounds, such as hormones, neurotransmitters, or certain drugs, to effector proteins. G proteins control the activity of hormone-sensitive adenylate cyclase in essentially all cells, the activity of a photon-activated cyclic GMP-specific phosphodiesterase in the retinal rod outer segment, and, presumably, the activity of other proteins as well. Because of their key role, they have been the focus of interest by several investigators.

Dr. Gilman determined the amino acid sequence of a region that appears to be common to the α subunits of all G proteins, as well as a portion of the *ras* oncogene product. He then used this amino acid sequence to produce an oligonucleotide probe (a 36mer with minimal redundancy), and in turn used this to screen a bovine brain cDNA library. The only clone that was obtained with this probe has a cDNA insert of approximately 1700 base pairs. Analysis of protein and DNA sequences indicates that the clone codes for a G protein which is distinct from both $G_{o\alpha}$, the most common G protein in bovine brain, and the α subunit of transducin. A peptide was synthesized according to the deduced amino acid sequence of the cDNA clone; antibodies prepared against the peptide react specifically with $G_{s\alpha}$, the G protein that stimulates adenylate cyclase activity. In addition, RNA that hybridizes with probes made from the clone is detectable in wild type murine S49 cells by Northern analysis. However, mutant (cyc^-) S49 cells, which are essentially devoid of $G_{s\alpha}$ activity, lack this mRNA. Thus, the cDNA clone probably corresponds to $G_{s\alpha}$. This work will greatly facilitate study of the synthesis and action of Gs under normal and pathological conditions.

A portion of Dr. Bourne's recent research has focused on transducin, a guanine nucleotide binding protein of retinal rod cells that mediates the activation of a cyclic GMP phosphodiesterase in response to photoexcitation of rhodopsin. This signal transducing protein is a heterotrimer (M_r: α = 39,000; β = 35,000; and γ = 8,000) that, as indicated above, exhibits striking structural and functional homologies with the stimulatory (G_s) and inhibitory (G_i) coupling proteins of hormone sensitive adenylate cyclase. Dr. Bourne's laboratory recently isolated and sequenced a cDNA from bovine retina that encodes the α subunit of transducin. The cDNA was isolated with an antibody probe from a bovine retinal cDNA library in *lambda* gt11. The 2.2 kb cDNA hybridized to a single mRNA species in extracts of bovine retina but not of other tissues.

The predicted coded sequence of the cDNA includes the amino acids that are ADP-ribosylated by cholera and pertussis toxins. It also includes four segments,

ranging from 11 to 19 residues in length, that exhibit significant homology to 4 regions found in several other GTP-binding proteins, including the elongation factors of ribosomal protein synthesis in bacteria, EF-G and EF-Tu, and the ras proto-oncogene products of man and yeast. Because previous genetic and biochemical evidence indicates that each of these four regions participates in binding or hydrolysis of GTP in the other proteins, the four homologous segments in the subunit of transducin presumably form part of the GTP binding site. The amino acid sequence deduced in this study, in combination with other studies of tryptic fragments, makes it possible to predict which portions of the polypeptide interact with other molecules involved in retinal phototransduction, including the $\beta\gamma$ component of transducin, rhodopsin, and phosphodiesterase.

Studies from other laboratories have shown homology between the carboxyl terminus of transducin's γ subunit and the carboxyl terminus of the ras proto-oncogene product, p21, comparable in degree to the homologies between amino terminal sequences of the GTP-binding domain of p21 and amino terminal sequences of the α subunit of transducin. Thus, both the ras and G/transducin protein families may derive from a common evolutionary precursor.

"Hepatic Microsomal Enzyme Metabolism of Anesthetics"
R01 GM 23029-07 (Kaminsky, L.), New York State Department of Health

Previous studies showed that 2,2,2-trifluoroethanol is the cytochrome P-450 catalyzed metabolite of the volatile anesthetic agent fluroxene and its ethyl and vinyl analogs, and that this metabolite is responsible for the lethality of the anesthetic agent in rats.

Recent studies by Dr. Laurence Kaminsky reveal a novel and highly complex mechanism whereby the metabolite, trifluoroethanol, expresses its lethal effects. Trifluoroethanol produces marked edema in the small intestine and ileocecal junction, and the associated lesions permit bacterial infection. It also markedly decreases the bone marrow-nucleated cellularity. As a consequence, the peripheral leukocyte count decreases to 15 percent of control values and infecting bacteria are able to proliferate. Trifluoroethanol appears to have a direct effect on bone marrow cells, as was demonstrated in in vitro studies with bone marrow cell suspensions. In addition, in vivo studies show that trifluoroethanol synergistically enhances the lethality of bacterial endotoxins by approximately 1000-fold. The clinical symptoms associated with death from trifluoroethanol which is generated metabolically from fluroxene are very similar to those associated with a lethal dose of endotoxin. Metabolically produced trifluoroethanol thus causes lesions which permit bacterial infection, destroys the host defenses against the bacteria, and synergistically enhances the lethal effects of the endotoxins. This complex mechanism of action for such a relatively simple molecule may provide insight into the toxic effects of many other halogenated xenobiotics.

"Control of Intracranial Pressure During Anesthesia"
R01 GM 33006-02 (Artru, A.), University of Washington

For patients with head trauma, hyperventilation is used to decrease elevated intracranial pressure, thereby minimizing the likelihood of brain damage. Patients with head trauma who are being treated by hyperventilation often require anesthesia for surgery to repair life-threatening effects of trauma.

Without anesthesia, the intracranial pressure-lowering effect of hyperventilation lessens over a period of hours and may be lost with time, so that patients again are at risk for brain damage. Dr. Alan Artru is determining whether commonly used anesthetics alter the changes in intracranial pressure that are induced by hyperventilation, using procedures that permit the simultaneous determination of intracranial pressure, cerebral blood volume, and cerebrospinal fluid volume.

With no potent anesthetic present and intracranial pressure initially normal, hyperventilation maintains intracranial pressure at low values over four hours. With fentanyl anesthesia and intracranial pressure initially elevated, hyperventilation causes a 50 percent reduction in intracranial pressure, which is sustained for 2.5 hours, with intracranial pressure increasing to a 30 percent reduction by four hours. No expansion of the cerebral blood volume over four hours was seen, but a slight expansion of cerebrospinal fluid volume was detectable after 2.5 hours. With halothane anesthesia and the intracranial pressure initially elevated, hyperventilation caused only a slight reduction of intracranial pressure, and by four hours the intracranial pressure was elevated by about 30 percent. Both cerebral blood volume and cerebrospinal fluid volumes expanded during the four hours of hyperventilation. In contrast, with no potent anesthetic present but with the intracranial pressure initially elevated by an intracranial mass, hyperventilation results in a 30 percent reduction in intracranial pressure which is sustained over four hours. The cerebral blood volume expanded by about 1 ml and the cerebrospinal fluid volume contracted by the same amount over four hours.

Thus, fentanyl slows the recurrence of elevated intracranial pressure during hyperventilation via cerebrovasoconstriction. In contrast, halothane augments the recurrence of elevated intracranial pressure during hyperventilation. The cerebral blood volume expansion in the presence of halothane was similar to that seen in the absence of a potent anesthetic, and the cerebrospinal fluid volume also was increased, indicating impaired reabsorption.

This work should demonstrate which anesthetic(s) are safest for patients being hyperventilated to control intracranial pressure and also should provide information as to the mechanisms that are responsible for the recurrence of elevated intracranial pressure.

"Anesthesia-Analgesia of Awake Cat Spinal Neurons"
R01 GM 29065-05 (Collins, J.), Yale University

Dr. J. Collins is seeking to explain how an anesthetic produces anesthesia. Toward this goal, he has developed a technique for the neurophysiological recording of extracellular single unit activity from neurons in the dorsal horn of the spinal cord and to evaluate the effects of anesthetic agents on spinal neuronal activity. He has used this technique to measure the spontaneous activity in low threshold neurons, i.e. those neurons responsible for signalling low intensity tactile information. In control, drug-free animals, no low threshold neurons had spontaneous firing rates greater than two impulses per second, and many of the neurons had spontaneous firing rates less than 0.5 impulses per second. In contrast, in pentobarbital anesthetized animals approximately 1/3 of the neurons recorded thus far have spontaneous firing rates greater than two impulses per second, with some neurons demonstrating rates of 10-15 impulses per second.

The absence of spontaneous activity in control animals suggests the existence of descending modulatory influences that inhibit spontaneous activity. The presence of spontaneous activity in low threshold neurons in anesthetized animals, and the knowledge that inhibitory systems may be more susceptible to anesthetic actions than excitatory systems, suggests that the spontaneous activity seen in anesthetized animals may be due to anesthetic inhibition of descending inhibitory systems. Release of spontaneous activity at the spinal level by barbiturates may provide a partial explanation for changes in electrical activity that are seen in supra-spinal sites during clinical anesthesia.

The signalling capability of the central nervous system is assumed to be based upon the presence of a signal imbedded in noise (spontaneous activity). Thus, the increased spontaneous activity that occurs in the presence of barbiturate anesthesia may produce a noise level which interferes with normal signal processing, thereby disrupting the ability of the central nervous system to communicate with the periphery.

"Anesthetic Reactions in Surgery"
RO1 GM 31382-02 (Louis, C.), University of Minnesota at Minneapolis-St. Paul

Malignant hyperthermia is thought to be triggered by an uncontrollable release of intracellular calcium in muscle, resulting in a massive increase in metabolic rate and muscle contracture. A defect in the ability of muscle surface membranes to control calcium movement could contribute to the onset of malignant hyperthermia. Dr. Charles Louis has prepared sealed vesicles from the muscle membranes of pigs that are susceptible to malignant hyperthermia and is determining whether their calcium permeability is abnormal. These vesicles contain an ATP-dependent calcium pump whose affinity for calcium appears capable of lowering sarcoplasmic calcium to the levels present in relaxed muscle. Furthermore, the activity of this pump was found to be regulated by both the calcium receptor calmodulin and the cyclic nucleotide receptor, cyclic AMP-dependent protein kinase. Thus, the transport of calcium across muscle surface membranes appears to be sensitive to the circulating levels of hormones. In addition to the ATP-dependent calcium transport process, the muscle surface membranes also contain a sodium-calcium exchange system that could be one of the multiple mechanisms by which the muscle surface membrane regulates intracellular calcium levels. As the sodium-calcium exchange process seems to be regulated by ATP, this could have implications in malignant hyperthermia where muscle ATP levels are depressed.

The mechanism regulating calcium release from the sarcoplasmic reticulum of malignant hyperthermia muscle was compared with that of control muscle. After passively loading the sarcoplasmic reticulum vesicles with calcium, calcium release can be induced from the vesicles by a sudden elevation in external calcium. This release of calcium is enhanced by a number of agents, including caffeine and ATP, and is inhibited by others, including $MgCl_2$ and ruthenium red. The calcium concentration at which this release is half-maximally activated is in the micromolar range, which is in the range of calcium concentrations observed during muscle contraction *in vivo*. For sarcoplasmic reticulum from animals exhibiting malignant hyperthermia, calcium release is triggered at a lower external calcium concentration, and the amount of calcium released is greater than that from sarcoplasmic reticulum of control animals. Thus, there seems to be a defect in the control of the calcium release process in pigs susceptible to malignant hyperthermia. Whether this defect is in the sarcoplasmic reticulum

or in the transverse tubules communicating to the sarcoplasmic reticulum is under study.

This research should provide a better understanding of muscle function in malignant hyperthermia which, in turn, should lead to improved procedures for its diagnosis and treatment.

PHARMACOLOGICAL SCIENCES PROGRAM

Professional Staff

Christine K. Carrico, Ph.D.	Director
James Gilliam, Jr., M.D., Ph.D.	Program Administrator
Carl A. Kuether, Ph.D.	Program Administrator
M. Janet Newburgh, Ph.D.	Program Administrator
Elizabeth O'Hern, Ph.D.	Program Administrator

Support Staff

Ruth A. Linn	Secretary to the Director
Patricia P. Driscoll	Grants Clerk
Mary Beth Gallagher	Grants Clerk
Mary Catherine Wancheck	Grants Clerk

Pharmacology Research Associates

Stephen Buck, Ph.D.
Alison Cole, Ph.D.
Gail Handelmann, Ph.D.
Dana Hilt, M.D.
Brian Huber, Ph.D.
Stephen Ikeda, M.D., Ph.D.
Robert Levine, Ph.D.
William Maixner, Ph.D., D.D.S.
Charles Mantione, Ph.D.
Suzanne Patton, Ph.D.
Herbert Pierson, Ph.D.
Mark Brann, Ph.D.
Stephen Korn, Ph.D.
Mary Loeken, Ph.D.
Daniel Marshak, Ph.D.
Nancy Ostrowski, Ph.D.
Laura Plunkett, Ph.D.
Matthew Sills, Ph.D.
Jacqueline Smith, Ph.D.
Barton Weick, D.V.M., Ph.D.
Howard Wheeling, Ph.D.
W. Scott Young, M.D., Ph.D.

BIOPHYSICS AND PHYSIOLOGICAL SCIENCES PROGRAM

SCOPE AND OBJECTIVES

The Biophysics and Physiological Sciences (BPS) Program was established in 1984, by the expansion of the existing Physiology and Biomedical Engineering Program to include aspects of biophysics previously distributed through other programs within the Institute. Part of the rationale for this reorganization was the rapid growth in recent years of research in the biophysical sciences, and the need to provide a focus within NIGMS for such research.

As the title suggests, this Program is somewhat heterogeneous. The Biophysics Section focuses on research which results from the application of physical principles to biological systems, as well as the translation of these efforts, through engineering, into practical applications. Instrument development is also supported by this Section. The research support in the Physiological Sciences Section is primarily in the area of trauma and burn studies. This research is directed towards an improved understanding of the total body response to trauma, and the molecular mechanisms underlying this response.

The systems studied by these two sections are clearly at opposite poles in terms of their complexity. However, in both cases the intent is the same, i.e. to provide fundamental information about the biological processes involved. In the Biophysics Section, the primary goal is an understanding of the structure of biological macromolecules. In the Physiological Sciences Section, a major goal of this basic research is an optimal treatment of burn and trauma injuries.

In addition to research, this Program supports training in related disciplines through a variety of mechanisms. These include predoctoral support through institutional training grants as well as individual postdoctoral fellowships. In addition, the NIGMS Medical Scientist Training Program (MSTP) is administered through the BPS Program.

RESEARCH OVERVIEW

1. BIOPHYSICS SECTION

This section includes a number of research interests of which the study of biological macromolecules is predominant. It has long been appreciated that a full understanding of biological systems at a molecular level requires a prior understanding of structure-function relationships in biological macromolecules, particularly of proteins and nucleic acids. Perhaps more than most areas of biology, advances in this realm have been closely linked to advances in technology. The application of X-ray crystallography to determining the structure of myoglobin and, later, of DNA, showed how much information of biological importance could be obtained from such structural studies. X-ray crystallography still provides the most detailed picture available of molecular structure; however, this technique has its limitations, in that it gives primarily a static representation of what are known to be dynamic structures, and is confined to the crystalline state. The modern investigator has available a wide array of tools for the determination of structure at

different levels of detail and under different conditions. This program supports structural studies using techniques such as Raman spectroscopy, fluorescence, EXAFS, and EPR, all of which have unique contributions to make to specific systems under specific conditions. Of all the various spectroscopic tools, nuclear magnetic resonance (NMR) has emerged as the most broadly applicable. Indeed, recent developments have allowed a resolution of structure using NMR which in some cases approaches that of X-ray techniques.

In conjunction with these tools, molecular genetics offers the possibility of providing tailor-made macromolecules by the application of site-specific mutagenesis. The availability of such molecules allows hypotheses of structure-function relationships to be tested in a way not previously possible. As an example, this has given research related to the study of protein-folding a new life.

At the other end of the spectrum, theoretical developments in molecular dynamics, energy minimization of molecular structure, and distance geometry calculations have given new vigor to models of macromolecular structure and provided new computational tools for the determination of structure.

Together with the studies on structure just described, this Program supports research on the instrumentation and methods development which are critical to the progress of the field. This Program also supports instrument development in cell biology and in separation science. This includes, for example, cell sorters and electron microscopy, while progress in mass spectrometry has been particularly noticeable. Among the developments in mass spectrometry of particular importance to biology is the extension of the high mass range into regions where biomolecules that were previously inaccessible can be detected. Additionally, new approaches to the analysis of poorly volatile molecules have been of importance.

In practice, the program has evolved in such a way that the emphasis is largely on the development of methods and instrumentation for molecular and cellular analysis. This is not due to any lack of interest in broader physiological questions, but rather is a consequence of the natural focus of engineering and physiological studies on particular organ systems, or on other structural units. Consequently, much of the work in this area is being supported by the categorical Institutes of NIH.

A small portion of this program supports algorithm development, modeling, simulations, and heuristics.

2. PHYSIOLOGICAL SCIENCES SECTION

Trauma, including burn injury, is the leading cause of death among people 1 to 44 years of age and the fourth leading cause of death in all age groups. Burn injury, which is a special form of trauma, hospitalizes 75,000 people each year. One-third of these burn victims are children under age 15. Approximately 12,000 burn victims die annually, making burn injury the third leading cause of accidental death in the United States.

In the immediate 24-hour period after extensive trauma or severe burn injury, a cascade of systemic responses occurs. When these changes are prolonged,

they often culminate in respiratory and cardiovascular collapse, metabolic dysfunction, kidney failure, and overwhelming infection, any of which may lead to death. Among surviving patients, scarring and distortions of muscle tissue called contractures can result in major disfigurement and physical disabilities, accompanied by severe psychological stress and the need for extensive rehabilitation.

The major problem facing scientists in the field of trauma and burn research is the absence of an adequate understanding of the mechanisms that underlie the complications of severe injury, including burns. Much more needs to be known about the basic molecular, biochemical, and physiologic changes caused by this type of injury.

Research support in this section concentrates on studies relating to trauma and burn. These range from the molecular level to effects on the whole organism. Investigators supported by this program are studying the role of fibronectin in relation to shock and trauma; the nature of wound healing and the effect of growth substances on such processes; and the immunologic and physiologic alterations following burn injuries. A major development, contributed to by several groups, has been the evolution of an artificial skin.

All of these, and other studies, have as their principal objective the enhancement of our knowledge of how the body responds to trauma. Given the nonspecific nature of the insult, localizing and refining the mechanisms which initiate and mediate the response is a formidable task, and one which generally requires a multidisciplinary approach. Consequently, this section supports a number of trauma and burn centers which are usually placed in the context of larger trauma treatment centers. In FY 85, the Institute funded nine highly specialized trauma and burn centers. Although progress has been slow, because of the complexity of the problem and the difficulty of defining appropriate model systems, notable advances have been made which have reduced morbidity and mortality significantly.

SHARED INSTRUMENTATION

As noted in the 1983 Annual Report of the Physiology and Biomedical Engineering Program, the NIGMS Shared Instrumentation Program is continuing, albeit at a greatly reduced level. The reason for this is not a lack of interest, but rather that the Division of Research Resources (DRR) is now the major focus within NIH for requests of this kind. Consequently, the NIGMS program now restricts itself to applications that do not fall within the scope of the DPR guidelines. In practice, this means requests for equipment in the price range $30-100,000. In the last fiscal year, four such awards were made.

SMALL BUSINESS INNOVATIVE RESEARCH (SBIR)

The SBIR Act of 1982 mandated that federal agencies supporting scientific research set aside a certain percentage of their R&D budgets for applications from small businesses. The awards were expected to be of a more applied nature, ultimately leading to commercial products or services. Within NIGMS, a major category of this type of application is in the area of instrument and methods development. This includes a broad array of bioengineering projects.

In 1985, the BPS Program received 102 such applications, and funded 22, out of a total of 153 received by the Institute. The total dollars awarded within the BPS Program was slightly over $2 million.

Among the applications funded are the development of fiber optics for the transmission of a CO_2 laser beam, the construction of new magnets for in vivo NMR, and the development of computerized methods for scanning and transmitting radiological data.

NEW INVESTIGATOR RESEARCH AWARDS

In addition to the individual research grants (R01), the program project (P01) and research center (P50) mechanisms, this Program also has available New Investigator Research Awards (R23). These latter are intended to aid individuals in their transition from M.D. and subsequent research training, to becoming independent investigators. The New Investigator Research Awards (NIRA) are limited to three years. In order to continue their research, the investigators must compete for a regular research project grant. During FY 85, there have been six applications for new awards of which three were funded. There have been 12 grants active during FY 85, one of which was completed during the year.

RESEARCH TRAINING

Predoctoral Training - Systems and Integrative Biology

The training program in Systems and Integrative Biology (SIB) plays an important role in the NIGMS training picture. As part of the BPS Program, it has provided predoctoral training in both physiology and biomedical engineering.

Programs are multidisciplinary in scope but may emphasize one or the other discipline. The Institute has been insistent that biomedical engineering programs provide a strong biological orientation so that trainees may fully enter the world of biomedical research. There has been a similar emphasis on appropriately quantitative training for the more physiologicaly oriented trainees trainees. The SIB training also has a place in the overall Institute training approach. The SIB training programs concentrate on biomedical problems above the cellular level, in many respects taking up where the Cell and Molecular training grants leave off. There is often a meshing of the neurobiological interests of many SIB preceptors with the interests of preceptors in the Pharmacological Sciences training program. Since several of the SIB programs have a neurobiological emphasis, appropriate for a program stressing integration, regulation and interactions, the trainees may continue in research areas not supported by NIGMS. This is consistent with the NIGMS intent to prepare basic researchers for all biomedical areas, including those relevant to the categorical institutes.

The physiological programs, which have been an important component of SIB grants, have shown significant changes over the past several years. The number of students applying to graduate programs appears to be declining. This decrease is not observed in all graduate physiology departments, however, and in several institutions the quality of the entering student body is as good as or better than in past years, despite a decrease in applications.

At the same time, neurobiology appears to attract some of the best graduate students in biomedicine. Students of highest quality are still seeking out all areas of neurobiology: developmental, behavioral, and central processing.

For example, three SIB programs at major universities have admitted a record number of graduate students in neurobiology in 1985, because of the large number of superbly qualified candidates. All three have found the number of exceptional applicants in other areas to be declining, and have decreased the number of admittees in other areas. This strengthening of neurobiology relative to some other disciplines is also evidenced by the increasing fraction of MSTP trainees doing research in the area. In many institutions, such as the University of California, San Francisco, Washington University and the University of Chicago, the SIB strengths in neurobiology aid the overall training environment in MSTP. Of course, in many other institutions, such as Yale, Harvard, the University of Pennsylvania, and Johns Hopkins, the SIB programs in other areas of physiology and biomedical engineering, complement and strengthen the MSTP programs there.

When the SIB program was established, some concern was expressed that students in the more physiological graduate programs might not be seeking and obtaining training which was sufficiently quantitative. These concerns are turning out to be largely unfounded. The pervasiveness of computers in modern laboratories has enabled researchers to use more sophisticated data gathering, modeling and simulations than ever before. The top-ranked graduate students actually lead the way in computer use and, far from avoiding quantitative approaches, are seeking them out. Modern students are leaders in the electronic revolution.

There has been some turnover of training grants in the past fiscal year. A long-term grant and first-time grant failed to be rated sufficiently well for renewal. At the same time, several new grants were funded. At both the University of California at San Diego, and at Washington University, previously funded training programs were reorganized and new proposals were submitted following lapses in support. Both these new programs are excellent, and better organized as training programs than the previously-funded grants. The other new grant is to a newly funded joint program in biomedical engineering combining the engineering training at the University of California, Berkeley, with the biomedical research opportunities at the University of California, San Francisco. There are few biomedically-oriented bioengineering programs, and it is doubly pleasing to see a promising new one established, and be able to offer NIGMS training support. Some new proposals, requesting funding in FY 86, have been submitted and initially reviewed. They do show considerable merit and promise. The number of new, high-quality proposals attests to the vigor of the SIB training program.

Training Grants in Trauma and Burn Research

Eight postdoctoral training grants in trauma and burn research have been active during the past year, one of them (at the University of California, San Francisco) in its final year. All but one of the 22 appointed trainees hold M.D. degrees; the exception, a Ph.D. recipient, was appointed only for six months to complete his training. All but one slot was filled. One new training grant (at the University of South Alabama, Mobile) was activated on July 1st, thus

making no change in the total number of grants. The number of trainee positions awarded for the following year has decreased to 20.

The basic research pursued by these trainees includes studies on fibronectin in relation to shock and trauma at SUNY, Albany (Dr. Saba, Program Director); studies on osteogenesis and scar formation at the Massachusetts General Hospital, Boston (Dr. Burke, Program Director); studies on the role of prostenoids in pulmonary response to injury at University of Washington, Seattle (Dr. Carrico, Program Director); studies on mechanisms of lung injury associated with trauma and sepsis at Brigham & Women's Hospital, Boston (Dr. Hechtman, Program Director); studies on the role of calcium as an insulin secretogogue at Washington University, St. Louis (Dr. Monafo, Program Director); studies on non-invasive methods for monitoring the healing of fibial fractures at Beth Israel Hospital, Boston (Dr. Silen, Program Director); and studies in Germany on the permeability of lung capillaries after trauma and sepsis in relation to the ARDS by a University of California, Davis trainee (Dr. Blaisdell, Program Director).

Most of the former trainees are continuing in academic research and teaching or completing residency training.

Combined M.D.-Ph.D. Training - Medical Scientist Training Program

The Medical Scientist Training Program (MSTP) supports highly motivated students having outstanding potential for a research and academic career with integrated scientific and medical training leading to the combined M.D.-Ph.D. degree. The program's goal is to prepare its graduates to function independently in both basic research and clinical investigation.

MSTP grants are made to universities and their medical schools. The 25 institutions that currently receive MSTP support are responsible for the operation of their individual programs, including the selection of students to be supported. Collectively, these 25 programs offer the 683 carefully selected trainees a choice of a wide range of graduate programs in the biological, chemical, physical and social sciences.

Nationally, approximately 100 medical schools offer opportunities for M.D.-Ph.D. studies. Over the past few years there has been a noticeable increase in the number of medical schools expressing an interest in participating in the MSTP. Several have reorganized their combined degree programs so as to provide a central focus/administration for operating the program and advising students. Many have included the participation of all of the relevant graduate training faculty and have changed their admission practices so that a student is admitted on a school-wide basis rather than by graduate department. Some have taken the additional step of funding their own programs. This increased interest has already resulted in a number of new applications requesting MSTP support. Overall, it is clear that some schools have taken significant steps to bolster the quality of their combined degree offerings and may successfully compete for MSTP funds in the near future.

The graduates of the MSTP continue to distinguish themselves, both individually and collectively. A preliminary study is being made of the 1174 trainees who entered the program in the period between FY 1964-1978. While this study is

not yet complete, the findings to date indicate that as of 1983-84 about 70% of the trainees achieved the M.D.-Ph.D. degree (during this period some programs did not require all students to pursue both M.D. and Ph.D. degree study); 5% are still working to complete their degree studies. Of these dual degree graduates, 431 are still in training as residents, fellows or research associates; 351 have completed their post-graduate studies and made career choices (88% are in academic or research positions, 12% are in medical practice.) It is of interest to note that of those who did not complete dual degree study but who have finished their post-graduate training, 60% hold academic or research positions.

RESEARCH HIGHLIGHTS

Biophysics

Among the methods for determining macromolecular structure, X-ray crystallography remains paramount. Consequently, developments in technology in that area are of great significance. These include data collection procedures and improvement of refinement methods. The highlights presented in this area utilize a variety of such new methods in studying systems as varied as lipids and protein-nucleic acid interactions.

The study of proteins involved in the regulation of gene expression is of major importance. Studies of the structural basis of such regulation are seen through examples not only from X-ray crystallography, but organic synthesis of synthetic restriction enzymes, and, in particular, the use of NMR.

New developments in NMR show considerable promise in being able to determine structures in aqueous solution at atomic resolution. This is shown through studies on both nucleic acids and proteins.

A variety of other techniques is also available to study molecular structure. Although these do not have the general applicability of X-ray crystallography and NMR, in particular circumstances they have unique value and can provide information not otherwise available. Examples are provided for Raman spectroscopy and electron microscopy.

Finally, this Program supports a variety of studies on modeling and mathematical methods. One such example, modeling of human circadian rhythms, is given.

"Structural Basis of Enzyme Action, Specificity, and Control"
Project #5 of 5 P01 GM22778-10 (Steitz, T.) Yale University

Dr. Steitz's X-ray crystallographic studies on the structure of proteins and nucleic acids have yielded important information on the mechanism and role of ligand induced conformational changes in protein and enzyme specificity and control. Previous work from his laboratory has focused on the nature, role and mechanism of conformational changes in the catalytic action of yeast hexokinase. His work on the *lac* repressor core and other DNA-binding proteins has led to new knowledge on protein-DNA recognition. His laboratory has an impressive ability to generate excellent and scientifically significant projects and to generate crystals of sufficient quality for structural determinations. The latest, and perhaps most exciting, of these projects involves

resolvase, the site-specific recombination enzyme from the prokaryotic transposable element $\gamma\delta$.

Resolvase is involved in the insertion and excision of transposable elements into the host DNA. The primary function of resolvase is to catalyze an intramolecular site-specific recombination between two identically oriented copies of $\gamma\delta$. This enzyme also serves as a transcriptional repressor that regulates expression of the divergently transcribed transposase and resolvase genes. During recombination (the breaking and rejoining of DNA strands to produce new combinations of DNA molecules), resolvase acts as a site-specific topoisomerase whose activity is triggered by interaction of two sites on the DNA which the resolvase recognizes and binds to. There are a total of three binding sites on the DNA, spanning about 120 base pairs of the intergenic region. Since all three sites are required for maximal recombination efficiency, they define the recombination site, res. The recombination reaction and site have previously been characterized, making this an interesting protein for structural studies.

The structural determination of resolvase is difficult since it is insoluble at low ionic strength and thus not easy to crystallize. Nevertheless, Dr. Steitz and his coworkers have obtained crystals of native resolvase and have determined a low resolution structure, showing the molecule to have two domains and to be a tetramer in the crystal. Since the crystals diffract poorly, it is unlikely that high resolution results will be obtained from these crystals. Fortunately, however, the protein can be proteolytically cleaved into two fragments, equivalent to its two domains: a large fragment that forms all of the subunit contacts and presumably catalyzes the recombination reaction, and a small fragment that binds specifically to the sites in the transposable element. The large fragment crystallizes in several well-ordered forms. Dr. Steitz is collecting high resolution data from these crystals on an X-ray area detector and is utilizing site directed mutagenesis to introduce a sulfhydryl group in the hope of producing a good heavy atom derivative and thus solve the phase problem inherent to crystallographic projects. Attempts to crystallize the small fragment on its own and as a complex with DNA fragments of about 10 base pairs also seem very promising. Dr. Steitz's structural studies should indicate the mechanism by which resolvase recognizes specific DNA sequences. Additional questions will focus on the role and function of the three different DNA binding sites, the mechanisms of recombination, and the mechanism of movement of the two DNA duplexes and protein subunits. Dr. Steitz has chosen an exciting and challenging problem, and is making impressive progress.

"X-Ray Diffraction Studies of Vertebrate Eye Membranes"
RO1 GM 32614-06 (Reynolds, G. and Gruner, S. M.) Princeton University

Biomembranes are amongst the most important, and least understood, parts of the living cell. A typical biomembrane consists of roughly equal masses of protein and a mix of lipids. The proteins act as the chemically specific machinery of the membrane, performing such tasks as transport, hormone recognition, and enzymatic catalysis. The lipids are a class of relatively similar molecules which aggregate to form the planar bilayer matrix in which the proteins are imbedded.

The properties required of the lipid molecules are poorly understood. To illustrate the problem, consider that the different kinds of biomembranes in a cell each have quite different lipid compositions. If, by reasons of diet or

mutation, the cell is denied a specific lipid component, the membranes perform a very complicated readjustment of their entire lipid spectrum in an attempt to compensate. Often the cells can adjust, resulting in an entirely different set of lipid compositions. The physical characteristics of the membrane which are being adjusted by this lipid turnover are not known. Our knowledge of the process is so limited that basic research is still struggling to find the right physical processes to explore and the right questions to ask. Yet the relations to health are immediate and profound. For instance, if mammals are fed a diet deficient in essential fatty acids, which are necessary precursors for synthesis of certain lipids, the membrane readjustment process fails and the animals die with multiple pathological symptoms.

The effects different lipids confer on the physical properties of bilayer membranes are being investigated by Drs. Sol M. Gruner and George T. Reynolds. These studies were undertaken in an attempt to understand the interplay of lipid composition and structure in vertebrate retinal membranes. In the last decade it has become known that roughly half the lipids of a typical biomembrane are of a type which spontaneously form bilayers when purified, whereas the remaining lipids (dubbed "non-bilayer" lipids) do not, by themselves, readily form bilayers. Moreover, the essential dietary fatty acids mentioned earlier are required for synthesis of the "non-bilayer" lipids. The route taken by Drs. Gruner and Reynolds has been to formulate a statistical mechanical theory of molecular characteristics which distinguish bilayer from "non-bilayer" lipids. Inherent in the theory are a set of measurable physical parameters whose values determine if a lipid mix is stable as a bilayer or a non-bilayer configuration. Most importantly, the theory yields direct insight as to how molecular features combine to result in final values of the physical parameters. The theory allows one to understand how, and if, a readjustment of the ratios of lipids in a mix may be performed to compensate for the loss of a component of the mix to allow the parameters to be kept constant.

Experimental verification of the theory has been obtained over the last two years by using the technique of x-ray diffraction to measure the structure of lipid phases. An important physical parameter of the theory was the intrinsic radius, R_o, through which a lipid layer must be bent to minimize its elastic energy. It was discovered that the addition of several percent dodecane to a lipid mixture allowed the system to form cylinders of lipid, the radii of which were at the elastic minimum. X-ray diffraction of the bilayer lipids dioleylphosphatidylcholine (PC) and dioleylphosphatidylethanolamine (PE) indicated that the former had a large value of R_o while the latter had a small value. Experiments performed on PE-PC mixtures verified that the R_o value of the mixture varied systematically with the PE fraction and correlated well with the minimum temperature at which a bilayer configuration was destabilized and transformed into a nonbilayer phase. Thus, it was shown that the intrinsic curvature, R_o, is a property of the lipid layer which measures the stability of the bilayer. Bilayers are unstable when R_o is small.

In collaboration with Drs. P. Cullis and C. Tilcock at the University of British Columbia, a combined x-ray diffraction, NMR, and electron microscopic study was undertaken to examine molecular features which affect the R_o value of lipids. The classic bilayer and "non-bilayer" lipids PE and PC, respectively differ only in the degree of methylation of the molecule. Consequently, lipids were synthesized with intermediate degrees of methylation. These were

then subjected to structural analysis to determine how methylation was affecting the intermolecular forces. The results of these studies demonstrated that water competes with inter-lipid hydrogen bonds in bonding to the lipid phosphate groups. The effect of methylation was to shift the competition in favor of the water, thereby lessening inter-lipid forces in the polar part of the lipid layer. The intrinsic curvature, R_o, is related to the ratio of the polar to non-polar forces; thus, the net effect of methylation was to increase R_o. The same effect has been recently measured in PE-PC mixtures in which the average number of methyl groups per molecule was adjusted by varying the PE to PC ratio. Another result of this study was insight as to why dehydration of a lipid system promotes the formation of non-bilayer phases: removal of surface water shifts the hydrogen bonding equilibrium in favor of inter-lipid bonds, thereby reducing R_o and destabilizing the bilayers.

The experiments and theory combine to yield detailed insight as to what distinguishes bilayer and non-bilayer lipids. Most fundamentally, each component contributes to the elastic bending energy via individual contributions to a final value of R_o which characterizes the lipid mix. Packing constraints compete against deformations of the lipid sheet by raising the local free energy. If R_o is small, then a tight local deformation can compete favorably against the packing constraints, which fight the deformation, and non-bilayer organizations result. If R_o is large, then bending deformations are energetically unfavorable, and bilayers are the stable structures. However, intermediate values of R_o, which are obtained when bilayer and "non-bilayer" lipids are mixed, form marginally stable bilayers which can be locally deformed by small perturbations of the packing contraints.

Research emphasis is now shifting to measurement of the parameters for various real biomembranes. Drs. Gruner and Reynolds have postulated that R_o is the critical parameter that is adjusted for each type of membrane and that the packing parameters are locally modulated by the protein mediated insertion of hydrophobic biopolymers. The fact that real biomembranes contain both bilayer and "non-bilayer" lipids suggests that the R_o value should result in marginally stable bilayers. This is precisely the regime over which local packing modulation is most effective in exerting local control of the molecular geometry. In other words, it may be that nature does not want extraordinarily stable (eg. large R_o) bilayers because there is no way to locally destabilize them. Rather, overall stability can be achieved via control of the intrinsic curvature parameter while allowing for the controlled local instability required for cellular functions such as endocytosis and membrane fusion. Further measurements are needed to test these ideas. However, if they are correct, they will provide a much needed rationale for the reasons behind the lipid composition of biomembranes.

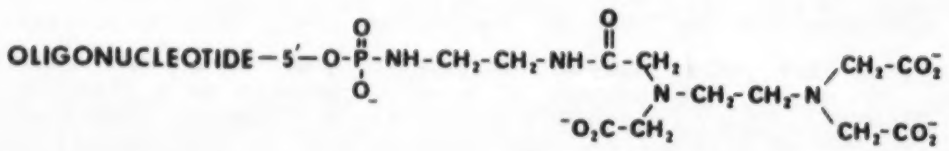
"A Synthetic Restriction Enzyme - A Step in the Right Direction"
R01 GM 33023-01 (Orgel, L.) Salk Institute for Biological Studies

Restriction enzymes are essential tools for modern molecular biologists. As a result of their ability to bind to (recognize) specific sites on DNA and cleave DNA at that site, they have become useful tools for such procedures as gene mapping and the replacement of specific sequences in genes. A limitation to the use of restriction enzymes has been that a scientist has to depend on the generosity of mother nature to produce a restriction with the desired specificity.

A major step in overcoming this difficulty has been taken by Leslie Orgel and Barbara Chu at Salk Institute in work supported by GM 33023. In this project they synthesized an organic molecule which cleaves single stranded DNA at a specific site. An important feature is that the specificity can be varied at will.

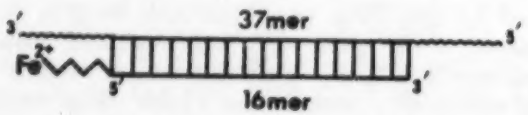
This project represents an extension of the elegant work of Peter Dervan and his colleagues at the California Institute of Technology (supported by GM 27681). In that work, Dervan synthesized compounds that mimic the cancer chemotherapeutic drug, bleomycin, in its mechanism of action. That is, his compounds contain a moiety with DNA binding properties (e.g. an intercalator such as daunomycin) and an iron(II)EDTA moiety. The iron(II)EDTA moiety proved to be the business end of the molecule in that in the presence of atmospheric oxygen and dithiothreitol, free radicals capable of cleaving DNA are produced. The DNA binding moiety acts to localize the destructive action of the radicals to a site on the DNA. Orgel and Chu have taken this project one step further. They have replaced the rather non-specific DNA binding portion of Dervan's compounds with an oligonucleotide of defined sequence. Ideally, this oligonucleotide should bind specifically to a complementary sequence on another strand of DNA and cleave the DNA at that site. The beauty of this procedure is that, theoretically, the binding site (and hence the cleavage site) can be specified by the sequence of the oligonucleotide binding portion. Given the ease with which specific sequences can be synthesized, it should be possible to synthesize an artificial restriction enzyme of any desired specificity.

Although this scheme works fine on paper, in practice Dr. Orgel's initial experiments have shown that there are still many obstacles which need to be overcome. Specifically, what Dr. Orgel has done is to synthesize a test DNA sequence 37 nucleotides long and a cleaving compound. The cleaving compound consists of a sequence 16 nucleotides long which is complementary to a portion of the 37-mer, and this 16-mer is attached via an ethylenediamine linker to an iron(II)EDTA moiety. The individual portions of the molecule are shown below:



binding portion linker EDTA

The idea behind the experiment is that the cleaving molecule binds to the test strand and cleaves at the site shown.



Under actual experimental conditions, the cleavage molecule is incubated with the test strand labeled with 31P at the 5' end, dithiothreitol, and poly-A which functions as a free radical trap. The reaction is stopped by addition of excess EDTA (to sequester the iron from the cleavage compound) and the results analyzed by gel electrophoresis. The major product was shown to be the expected 26-mer with breakage at adjacent residues. A minor product was a 16-18-mer produced presumably by the iron(II)EDTA moiety doubling back along the chain as a result of the flexibility of the ethlyenediamine linker.

Thus, although initial results are very encouraging, there are still problems to be overcome. The first is the less than perfect specificity of the cleavage reaction. Orgel and Chu suggest that a less flexible linker might be used. The second difficulty is that the yield of the reaction is low (only about 15% of the test strand is cleaved). Adjusting conditions may serve to optimize the yield. Another approach may be to select a cleaving moiety more potent than iron(II)EDTA. Finally, the major limitation is that, as formulated, the "artificial enzyme" cleaves only single-stranded nucleic acids because it does not have sufficient affinity to displace the complementary strand from double stranded DNA. While these compounds might exhibit great utility in, for example, mapping single stranded viral DNA or RNA, the utility of restriction enzymes has been that they cleave double stranded DNA. Hence, future work will need to focus on increasing the affinity of these molecules.

Hidden in this initial success is that Chu and Orgel's real contribution has been the development of the chemistry which allows almost any organic group to be attached to the 5' end of an oligonucleotide. Thus, while an obvious application is that described here, there are others, such as attaching signals to probes. In the end, what Chu and Orgel envision is that someday we might know what specific sequences cause genes to become oncogenes and lead to cancer. It then might be possible to design a compound with built in specificity to destroy oncogenes and leave normal DNA intact.

"Studies of Protein Internal Motion by NMR"
R01 GM 23289-08 (Griffin, R. G.) Massachusetts Institute of Technology
R01 GM 24266-07 (Opella, S. J.) University of Pennsylvania

Over the past five years, interest in protein dynamics has been rapidly increasing. Of course, the concept of a native protein as a flexible rather than a rigid structure is of much earlier origin. However, both theoretical and experimental studies provide new evidence that is considerably more detailed in its description of molecular motion. In particular, nuclear magnetic resonance (NMR) has been of great value. The most common application of NMR is to the study of molecules in solution. However, orientation dependent NMR-parameters get averaged out by the rapid rotational and translational movement of the protein. In contrast, measurements in the solid state allow the observation of internal motions over a wide range of time scales.

Model studies of crystalline amino acids and peptides show not only amino and methyl group rotations, but also significant side chain rotation. Crystalline phenylalanine deuterated in the five position has an NMR frequency that depends only on the angle between C-^{2}H axis and the imposed magnetic field. It is therefore easy to calculate spectra to conform to a number of situations, e.g. where there is no ring movement, where the rings flip between two equilibrium orientations, etc. Griffin and others have shown that the phenylalanine system

exhibits fast ring flips that are consistent with a two-site jump of 180°. Griffin has suggested that L-[d5] phenylalanine contains two different crystalline forms, one of which permits fast ring flips. Studies with other amino acids show an inconsistent pattern, with some exhibiting motion and others not. It appears that the internal dynamics are very sensitive to crystal packing. A correlation between X-ray crystal structure and NMR results should provide useful information on this point.

Similar studies have been performed by Opella and his colleagues on aromatic rings in the coat protein of the filamentous virus fd. This system is examined in solutions, since mobility of the structure is so low, and thus solid state methods must be used. Opella has observed clear motional patterns for phenylalanine and tyrosine, with jumps almost identical to the 180° flips seen in the crystalline phenylalanine. Conversely, the single tryptophan in the coat protein does not show such movement, although it is not clear that mobility can be entirely ruled out.

These and other results provide an experimental counterpart to the theoretical studies of internal dynamics in proteins. Further, this is an approach that cannot be duplicated by such high-resolution but static methods as X-ray crystallography. With improved methods of data acquisition and analysis, important advances should be forthcoming.

"The Structure of Actin Determined from Electron Micrographs"
R01 GM 26723-06 (Smith, P. R.) New York University Medical Center

The field of structural biology ultimately attempts to reveal biological structure at atomic resolution. Until relatively recently it has been assumed that X-ray crystallography was the only method which could generate the three-dimensional molecular structure of macromolecules at this level of detail. However, the studies of a number of investigators, (in particular the laboratories of Klug and of Unwin and Henderson), have demonstrated the power of image processing techniques in obtaining three-dimensional structural information from electron micrograph images. The importance of this should not be underestimated, since many biological structures cannot meet the rigorous criteria of crystallization of isomorphous replacement required for high-resolution X-ray analysis.

There are, however, a number of factors preventing the direct visualization of high-resolution structural details in proteins by electron microscopy. These include specimen preparation artifacts, radiation damage, the low contrast (and thus low signal-to-noise) of biological materials, and the loss of three-dimensional information arising from the great depth of field of a transmission electron microscope.

However, these contraints can to some considerable extent be overcome by image enhancement and three-dimensional reconstruction of low-dose electron microscope images. Image enhancement consists primarily of improving the signal-to-noise ratio of the structure of interest. This can be done either by superimposing a set of images representing common views of a material, or by obtaining the structure in the form of an ordered array. The material can then be averaged over the array using Fourier transform methods together with spatial filtering. The latter is by far the method of choice, if appropriate ordered arrays can be obtained.

The need for three-dimensional reconstruction arises from the fact that electron microscopes have a depth of focus of several hundred angstroms. Consequently, a micrograph is a projection image in two-dimensions of structural features from different depths in the specimen plane. The reconstruction once again involves Fourier transform techniques, and the use of a set of two-dimensional projection images taken from different directions. In the case of a crystalline sheet, this involves tilting the sheet with respect to the electron beam.

Dr. P. R. Smith, of New York University, in collaboration with Drs. Polland, Aebi, and Fowler at Johns Hopkins, has attempted to determine the structure of actin from Acanthamoeba by the use of image reconstruction techniques. A detailed understanding of actin polymerization and of actin interaction with myosin and other proteins has been limited by a lack of detailed structural information about both the actin molecule and the actin filament. Smith and his colleagues have attempted to determine actin structure from rectangular actin sheets with highly ordered areas. The sheets are negatively stained to help stabilize the structure and provide contrast. The crystalline sheet was tilted at different angles up to 60 degrees to provide a set of projection images. Three-dimensional reconstruction yields an actin molecule which appears as an ellipsoid of dimensions 56Ax33Ax43A. In more detail, it appears to be wedge-shaped, with a base of 43A and a top of about 25A. There is some indication that it is bilobed, with a shallow groove in the middle of the molecule separating the large base from the smaller top.

Paracrystals of the actin filament were also studied. These contain structural data to the 30A, compared to the approximately 15A resolution obtainable from the crystalline sheets. The filament model obtained from reconstructions indicates the presence of wedge-shaped subunits with a length of about 55A along the filament axis. Although no available filament maps provide enough resolution to unambiguously trace the individual actin subunits, there are features which can be used as constraints in model building. Using these constraints and the actin model obtained from the crystalline sheets, a filament model can be built which satisfies the experimental data available from the filament. The results obtained are in good qualitative agreement with reconstructions obtained in other laboratories. The model shows wedge-shaped actin subunits aligned with their long axis approximately parallel to the filament axis, with the wedge pointing towards the "barbed" end of the filament. The filament thus has a distinct polarity.

It is clear that three-dimensional reconstruction of electron microscope images has many pitfalls and, in general, provides substantially poorer resolution then single-crystal X-ray studies. Nonetheless, it can provide information on systems such as actin, where X-ray data is limited, particularly for organized structures such as the actin filament. In addition, the use of actin sheets offers the possibility of looking at the interaction of actin and actin-binding proteins which may bind to the sheet. In order to improve resolution of structure details, the major limitations to be overcome are in the area of specimen preparation. In particular, the use of negative staining automatically limits the obtainable resolution of structural details to 15 to 30A. Attempts to improve both specimen preparation and methods for three-dimensional reconstruction are being made.

"Dynamic Aspects of DNA Recognition"
P01 GM 32614-02 (Reid, B.) University of Washington

Dr. Brian Reid and his colleagues in the program project on DNA structure and recognition at the University of Washington are closing in on the problem of how gene expression is controlled at the level of molecular and atomic detail. Their success is due to the development of two-dimensional nuclear magnetic resonance (2DNMR) methods to assign almost all of the hundreds of hydrogen atoms in the NMR spectra of repressor proteins and regulatory DNA sequences such as operators and promoters. They can thereby study the structure, dynamics and interactions of these molecules in solution. The DNA sequences used in their studies were synthesized in quantitites sufficient for NMR experiments in the program project core laboratory by Dr. Shan-Ho Chou.

Using 2DNMR techniques, Dr. Reid and colleagues in the program project (Drs. Gary Drobny, David Wemmer and J. Michael Schurr) initially mapped out all of the proton-proton networks between base pairs in the double helical DNA molecule, d(CGCGAATTCGCG). These studies were then extended, first to synthesize both strands of a DNA duplex containing the Pribnow -10 promotor TATAATG and then to assign the protons in the sequence of 2DNMR. Promotors are short regulatory DNA sequences that function in the control of DNA transcription by binding RNA polymerase such that RNA synthesis is correctly initiated. The Pribnow promotor sequence, or "TATA" box, is the one located most proximal to the site of initiation and is upstream at about -10 from the first transcribed residue. Dr. Reid and coworkers examined the prototype or consensus sequence

CGTTATAATGCG
GCAATATTACGC

in which the promotor is flanked to two CG pairs to stabilize the ends of the helix. In the above sequence, the central ten base pairs are the actual promotor sequence of the spc ribosomal operon and the central eight pairs are the sequence for a promotor in the SV40 virus. The imino protons and almost all of the nonexchangeable base and sugar protons were assigned and the conformation and stability of the molecule were studied. The helix assumes as fairly uniform B conformation and exhibits asymmetrical melting; it is relatively unstable in the center of the AT-rich region, despite being surrounded by GC base pairs. Similar studies have been carried out on the phage lambda OR3 operator DNA and vC3 single base-pair mutant.

In an unexpected sideline to this research, Dr. Wemmer discovered that interchanging the fifth and eighth base pairs of the Eco RI sequence CGCGAATTCGCG to form instead CGCGTATACGCG leads to extensive formation of a looped hairpin DNA structure in solution via internal "bulging" and cruciform propagation under physiological conditions. The formation of such hairpin structures in DNA was previously thought to require torsional superhelical twisting.

In parallel NMR studies, Dr. Reid and coworkers are also investigating repressor proteins. They are using strategies similar to those developed by Drs. Wuthrich, Ernst and colleagues in Zurich. By means of 2DNMR methods, Mr. Paul Weber, a graduate student in Dr. Reid's laboratory, was successful in sequentially connecting and assigning protons in adjacent amino acid residues throughout the lamda phage cro repressor protein. This represents the first complete assignment of a repressor protein NMR spectrum. It should pave the

way for future NMR experiments on repressor-operator complexes.

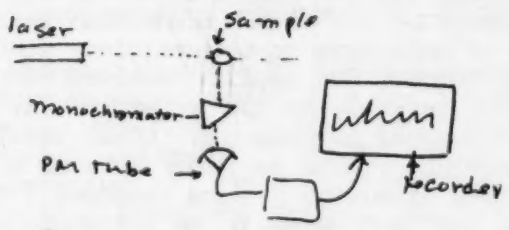
An even more interesting recent development is the ability to determine the three-dimensional structure of nucleic acids and proteins directly from solution using 2DNMR data combined with computer methods. The off-diagonal cross peaks in a two-dimensional nuclear Overhauser effect spectroscopy (NOESY) experiment identify pairs of protons that are sufficiently close (less than 5 Angstroms) to be dipolar coupled through space, provided the mixing time during which proton-proton cross-relaxation occurs is kept reasonably short. From the time-dependent intensities of the NOESY peaks, these distances can be calculated with a precision of 0.3 Angstroms. The resulting network of about 1000 distances can be thought of as threads of string tying down hundreds of known points in the molecule. A three-dimensional structure that conforms to all these constraints was calculated by means of matrix computer methods. Dr. Dennis Hare, a consultant to the program project and a former graduate student of Dr. Reid's, has written (with partial support from an NIH SBIR grant, GM 35620) an efficient computer algorithm for embedding the hundreds of experimental constraints from distance space into Cartesian 3-space. A simpler two-dimensional analogy of this problem would be the reconstruction of a map of the United States given only the distances between pairs of cities that are within 500 miles of each other. Using NOESY data alone, the three-dimensional structure of the DNA-recognition domain of cro repressor in solution has recently been determined; it agrees very well with the X-ray structure of this protein determined by Dr. Brian Matthews and coworkers at the University of Oregon (GM 20066).

With minor modifications, the computer program is equally capable of determining nucleic acid structure. Drs. Hare and Reid have now determined the complete solution structure of a DNA hairpin entirely from NMR data. This is the first such atypical non-duplex DNA structure to be determined at atomic resolution by any technique. From the progress of the program project, it appears that NMR is becoming an extremely important new tool for examining the molecular structure of important biological polymers with a level of precision rivaling that of crystallography, without the often frustrating requirement of growing crystals. Progress in these areas should now be accelerated by the availability of the new 500 MHz NMR spectrometer recently built by Dr. Gary Drobny, also with support from this program project grant. The only limit to the method is spectral resolution and, at current NMR frequencies of 500 MHz, polymers larger than 10,000 to 15,000 molecular weight may unfortunately be too large for structure determination.

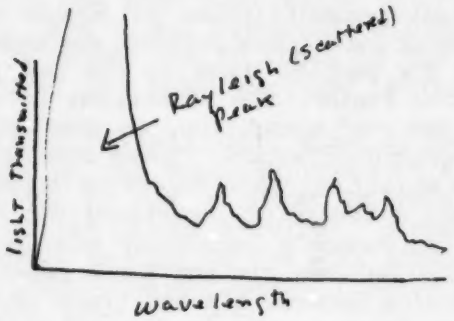
"UV Resonance Raman Studies of Protein Structure"
R01 GM 30741-04 (Asher, S.) University of Pittsburgh
"Ultraviolet Laster Resonance Raman Studies of Proteins"
R01 GM 32323-03 (Hudson, B.) University of Oregon

Although C. K. Raman discovered the Raman effect in the 1920's, a discovery for which he received the Nobel Prize in 1930, the use of Raman spectroscopy to study biological systems is relatively new. In its simplest terms, Raman spectroscopy makes use of scattered light to determine vibrational bands of a molecule. A diagram of a simple Raman spectrometer is shown below. The Raman bands are those which appear at lower energies from the Rayleigh or main scattering peak. The differences in wavelength represent vibrational frequencies

of the bonds being examined. These energies give fundamental information about the bending and stretching about these bonds. Raman spectroscopy has certain advantages over infrared spectroscopy, which measures essentially the same thing. The main advantage for biological systems is that vibrational spectra can be easily measured in water, a task which is difficult in infrared spectroscopy.



Simple Raman Spectrometer



Typical Raman Spectrum

For many years, however, the use of Raman spectroscopy remained dormant for lack of a suitably intense monochromatic light source. The situation changed dramatically with the introduction of lasers. A second major advance was the discovery that if a molecule is excited at an absorption band, the Raman effect is markedly enhanced leading to much greater sensitivity. The latter discovery, called resonance Raman spectroscopy, has led to the study of a number of proteins with prosthetic groups, e.g., heme proteins, which absorb light in the visible region of the spectrum. Although the Raman spectra of biological macromolecules are extremely sensitive to changes in secondary structure because vibrational bands are affected by hydrogen bonds involved in helical and sheet-structures, the use of Raman spectroscopy to probe protein and nucleic acid structures has lagged because, in the absence of a prosthetic group, these groups only absorb light in the ultraviolet (uv) region of the spectrum. Thus, in the absence of lasers which emit light in the uv, it was impossible to take advantage of the resonance effect.

The situation has changed dramatically with the introduction of pulsed Nd YAG lasers. Two NIGMS grantees, Drs. Hudson and Asher, have used various harmonics of the Nd YAG laser to develop instrumentation capable of exciting proteins and nucleic acids in the uv. Results from these laboratories, while still in preliminary stages, demonstrate the potential of this method.

Dr. Asher has taken advantage of the enhanced sensitivity of resonance Raman to study the Raman spectra of polycyclic aromatic hydrocarbons (PAH's) selectively in rat liver microsomes, a first step in a new program to examine PAH metabolism in an almost "in vivo" system. As a side product to this research, Dr. Asher developed a new optical dispersion device which can be tuned to "reflect" a narrow wavelength while transmitting light in adjacent spectral regions. The efficiency and narrow bandwidth of this device permit unique applications in spectroscopy. Importantly, Dr. Asher has constructed a continuously tunable uv Raman spectrometer, which may well serve as the basis for a low-cost high sensitivity commercial instrument.

The few studies which have been done on proteins and related models point to intriguing possibilities for future use. In complementary studies, Drs. Hudson and Asher have measured the resonance Raman spectra of N-methylacetamide, a model for peptide bonds in the far uv. These studies show that in the 190 nm spectral region, the transition is dominated by the pi-pi* transition rather than the n-p* transition, an observation of importance to theoreticians studying protein structure. These results also suggest that peptide backbone vibrations will also show selective enhancement in proteins for uv excitation at 190 nm, a result of importance for future studies on proteins. Other results from Dr. Hudson's laboratory show considerable promise in being able to distinguish between cis and trans peptide bonds involving proline residues. The conversion between cis and trans prolines has been shown to be extremely important in our understanding of the mechanisms of protein folding. Thus the potential ability to distinguish them spectroscopically has considerable potential for the study of protein folding.

Thus, although Raman spectroscopy has been used for over 50 years, it has been the development of new instrumentation in the last two years which point to the future use of Raman spectroscopy as a routine tool for the study of protein structure.

"Mathematical Modeling of Human Circadian Rhythms"
R01 GM 30719-01 (Kronauer, R.) Harvard College

Chronobiologists are striving to gain an understanding of a number of biological rhythms and the mechanisms (biological clocks or pacemakers) which control them. The biological events which have a near 24 hour period (commonly referred to as circadian rhythms) include core and skin temperature, cortisol secretion, REM (Rapid Eye Movement) sleep, and sleep-wake (SW) cycle. Experiments to date have determined that these circadian rhythms are controlled by two separate pacemakers which may or may not have a common neural substrate. The pacemakers differ in that one is a "deep" stable circadian oscillator which is not readily influenced by external factors. This "deep" oscillator controls the core temperature, cortisol secretion and REM sleep. It is known as the "x" oscillator, and it does not vary by much over an hour in either direction of its 24 hour period. The second circadian oscillator is more labile. It controls the skin temperature and the SW cycle. In free-run (desynchronized) individuals the labile or "y" oscillator may have a period of from 24 to 50 hours. Desynchronization can arise from either the lack of external time cues or from organic causes and can have a devastating effect on the subject.

Dr. Richard E. Kronauer, the Gordon McKay Professor of Mechanical Engineering

at Harvard University (GM 30719-01), developed a mathematical model of circadian rhythms with data obtained from a number of investigators at Harvard and Cornell. The model has been used to interpret data concerning entrainment of the "x" and "y" oscillators to external time cues ("Zeitgebers" or time setters) of varying lengths. The adjusted models were used to make predictions concerning these biological clocks with different time cues. The predictions of the model are validated by experimental results. Dr. Kronauer's model of two loosely coupled oscillators to simulate "x" and "y" pacemakers has successfully simulated a number of known experimental results. The model is now being used to design new experiments that should help to uncover the basic mechanisms of these biological rhythms and their physiology.

The model has contributed the discovery of a 3-4 hour zone during which the onset of sleep is inhibited (forbidden). This sleep inhibitory effect is controlled by the "x" oscillator and can influence the SW cycle of the "y" oscillator causing episodes of insomnia even in sleep deprived subjects.

Dr. Kronauer proposes to continue to modify this two oscillator model to better conform to experimental data and continue to use the validated model to simulate new experimental condition. He proposes to use a number of physical factors to determine if the "x" or "y" oscillators can be entrained. A greater understanding of circadian rhythms is necessary to improve the work schedules of airline crews and to help shift workers and transmeridional travelers to adjust to dramatic changes in their time cues.

Physiological Sciences

I. Inflammatory Response to Injury

Sepsis, shock, and pulmonary complications remain a major complication and cause of death in seriously burned patients. Definition of the immunologic and physiologic alterations following burn injury has been an active pursuit of many investigators. In recent years, many of these changes have been shown to be associated with the release of potent mediators which are actively involved in the tissue damage, particularly in the lungs of the burn victim. Understanding the biochemical interactions of important pathways, primarily complement and arachidonic activation and the formation of oxygen radical products, are prerequisite to developing protocols for the regulation of the inflammatory responses. It is now being demonstrated, as noted in the investigations of Dr. Peter Ward outlined below, that complement derived chemotactic substances, primarily C5a, are instrumental in the production of secondary lung tissue damage by the activation of neutrophils.

"Thermal Injury, Complement and Leukocyte Dysfunction"
R01 GM 28499-05 (Ward, P. A.) University of Michigan at Ann Arbor

These studies of skin thermal injury in rats have demonstrated the onset of the systemic activation of complement, acquisition of phagocytic cell defects, and the development of secondary lung injury.

Dr. Ward has shown that thermal injury to the skin of anesthetized rats results in the progressive consumptive depletion of total hemolytic serum complement activity and the appearance of a short-lived neutrophil chemotactic activity in the serum. The observation that the addition of antibody directed against human

C5a completely suppresses the chemotactic activity of burn serum in vitro, suggests that thermal injury of skin results in complement activation and generation of C5-derived chemotactic activity. Coincidental with the appearance of peak levels in serum of the C5-derived chemotactic activity is a transient neutropenia that may be caused by a combination of increased adherence of neutrophils to vascular endothelial cells and the sequestration of blood neutrophils within the interstitial capillaries of the lungs and other organs. The role of complement activation and sequestration of neutrophils in the development of acute lung injury, as demonstrated by significant increases in lung vascular permeability, were related by the finding of morphological changes postburn that included alveolar hemorrhage, interstitial edema and aggregates of neutrophils in pulmonary capillaries, and the observation that depletion of the complement system or blood neutrophils prior to thermal injury resulted in almost complete protection from acute lung microvascular injury.

The data obtained thus far suggest that lung vascular injury following skin thermal injury is related to the generation of oxygen-derived free radicals by complement-activated neutrophils. Treatment of thermally injured animals with catalase, an enzyme which converts hydrogen peroxide to oxygen and water, results in a dramatic reduction in lung damage. Additional observations indicate that the hydroxyl radical, an iron-mediated conversion product of hydrogen peroxide, may be the most important oxidant involved. Treatment with scavengers of the hydroxyl radical (dimethyl sulfoxide, dimethyl thiourea, sodium benzoate) or with iron-chelating substances (deferoxamine mesylate, 2,3-dehydroxybenzoic acid) protected the lungs of the thermally injured rat as assessed by changes in vascular permeability and by morphological parameters.

It has also been found that the same interventions (administration of catalase, hydroxyl radical scavengers and iron chelators) that protect against the accumulation of neutrophils in the lungs and the stimulation of these cells to produce lung tissue-damaging oxygen radicals also significantly diminish the tissue and plasma levels of lipid peroxidation products that appear after burn injury. There is an increasing body of evidence that these peroxidation products, conjugated dienes, may play a role in the suppression of phagocyte function. Overall, these experimental findings provide a basis for a better understanding of the pathogenic mechanisms leading to the pulmonary complication of burn injury and may also provide some suggestions for new therapeutic interventions in the treatment of burn victims.

II. Artificial Skin

A notable development has been the increased research interest and work on artificial skin. Following a severe and extensive burn, one of the prime needs is to quickly replace the skin, in order to avoid infection, prevent fluid loss, and retain body heat. The first two, especially, are life threatening problems that must be dealt with quickly, preferably in a matter of days. In the case of very extensive burns, the time honored practice of grafting the patient's own skin onto the wounds cannot be followed since an insufficient amount of unburned skin is available; other covering material must be found. Here the problem is the body's normal reaction to foreign material, be it synthetic or from another organism. Graft rejection is a familiar problem from transplant surgery and has been dealt with in a number of ways, as reflected in the examples given below. One way around the problem may be to amplify

the patient's skin in cell or tissue culture, either in vitro or in place on the patient. An obvious challenge, when these methods are used, is to minimize the time between the accident and the covering of the wound. Another desideratum is to minimize the pain and need for anesthesia that go with each grafting operation.

"Design and Evaluation of an Artifical Skin"
R01 GM 23946-09 (Yannas, I. V.) Massachusetts Institute of Technology

Dr. Yannas and Dr. Burke of the Massachusetts General Hospital have produced a biodegradable collagen-glycosaminoglycan membrane that functions as a dermal layer with a Silastic overlay. This material has been used on patients but has required final autografting when large areas have to be covered to provide an epidermal covering. In the second stage of this work, a small area of unburned epidermis is harvested from the patient at the time of the original excision of burned, dead tissue and processed to obtain a suspension of single basal cells. The basal cells are then seeded on the artificial dermis just under the Silastic sheet to form a new epidermal layer derived from the patient's own cells and then used immediately to cover the burned areas. This process has been accomplished using animals and is now being tested in patients.

"Burn Trauma Center"
P50 GM 21700-11 (Burke, J. F. and Ehrlich, H. P.) Massachusetts General Hospital

Dr. Ehrlich's wound healing project, a part of Dr. Burke's research center grant, includes a study of the modification of the present artificial membrane used by Drs. Burke and Yannas. The growth of new blood vessels into healing wounds is important for the generation of a new connective tissue matrix essential to the repair process. Using a wound healing model in rats and mice, Ehrlich has found that when the artificial collagen-GAG-Silastic membrane of Yannas and Burke is suspended in a heparin solution before its use as a graft, vessel growth into the synthetic material occurred within 7 days where none occurred before. This growth was associated with increased connective tissue deposition and cellular density of the graft.

On the other hand, the suspension of the artificial membrane in a solution of protamine prior to grafting inhibits the proliferation of fibroblasts, the deposit of connective tissue, and vascular ingrowth in the artificial membrane matrix. Thus, it may be possible to produce modified artificial skin where the quantity and quality of the new connective tissue formed can be controlled. The end results could be improved healing and better cosmetic results.

"Living Prosthetic Skin"
R01 GM 35068-01 (Hansbrough, J. F.) University of California, San Diego

Drs. Hansbrough and Boyce have developed methods for the convenient and rapid growth of human epidermal keratinocytes in defined media, and for the control of keratinocyte differentiaton in culture. They have also studied conditions for attachment and proliferation of keratinocytes on the collagen-GAG membrane produced by Drs. Yannas and Burke. Single cell suspensions of keratinocytes

are inoculated at a density of 1250 cells/cm^2 into a plastic petri dish containing the membrane and a culture medium supplemented with epidermal growth factor, insulin, hydrocortisone, ethanolamine and phosphoethanolamine. The investigators have defined various media requirements which will control the development of single cell line cultures when mixed cell populations containing both fibroblasts and keratinocytes are added to the culture media. They have found that discrete colonies of keratinocytes form on the membrane in 15-26 days and their growth is enhanced by the addition of human serum albumin and increased extracellular calcium ion concentration which encourages multi-layering of keratinocytes on the membrane. The porous structure of the membrane, however, permits it to be invaded by growing colonies of keratinocytes which is an undesirable characteristic. The current studies include modifying the pore structure at one surface of the membrane so that the keratinocytes will be restricted from penetrating the "external" face of the membrane while leaving the pores of the "internal" face large enough for fibroblast and vascular invasion. They anticipate that the procedure will allow cells from a small amount of the patient's unburned skin to be expanded rapidly and used to make a permanent graft with minimal risk of rejection.

"Composite Grafts in the Treatment of Thermal Injury"
R01 GM 30407-03 (Baxter, C. R.) University of Texas Health Science Center, Dallas

Dr. Baxter has attempted to develop a composite of cadaver skin dermis and vacuum blister-prepared sheets of autologous epidermis. The graft was originally tested on a rat burn model and successful engraftment with complete epidermal coverage and an absence of wound contracture was observed. Subsequently split thickness cadaver allografts (consisting of the superficial part of the dermis and the epidermis) were placed on full thickness burn wounds of eight patients. Three to five days later the entire epidermis was removed from the allograft and vacuum blister sheets of autologous epidermis were grafted to the exposed dermal surface. Grafting was successful in all patients. Follow-up studies of the wound areas have shown repigmentation of the grafted area with little or no contracture and minimal visible scarring. There has been no delayed graft rejection. The epidermal donor areas all healed without scarring and could not be differentiated from normal skin within four to six months.

"Burn and Trauma Research Center"
P50 GM 26145-07 (Shires, G. T. and Hefton, J. M.) Cornell University Medical Center

Dr. Hefton, as part of the studies being supported in Dr. Shires' Burn and Trauma Research Center, is developing a skin equivalent, composed of allogeneic epidermal cells cultured on a substitute dermis of collagen, which will provide a permanent covering for full-thickness burns. Partial-thickness pieces of cadaver skin were collected in sterile culture medium with fetal bovine serum and antibiotics. The epidermal layer was removed from the dermal part after incubation with trypsin and the epidermal sheets mechanically divided into single cell suspensions. The number and viability of the cells were determined and they were seeded into plastic culture flasks. Epidermal cells, grown in this manner, stratified in culture and yielded multilayered membranes of partially differentiated cells resembling skin epidermis. By examining the growth

characteristics of these cultured epidermal cells, it has been found the addition of calcium and epidermal growth factor (EGF) alters the cell morphology to a spindle-shaped cell as compared to the polygonal shape observed without EGF. In addition, the cultures reached confluency sooner and only stratified to a maximum of two to three cell layers. The addition of dexamethasone increases the attachment and proliferation of the cells and has allowed the investigators to increase the number of epidermal cells through serial passage. Further studies are underway using cryopreserved and thawed cells.

"Cyclosporin A and Skin Allografts for Burn Management"
R01 GM 31974-02 (Achauer, B. M.) University of California, Irvine

Cyclosporin A (CyA), a new immunosuppressive drug inhibits T-lymphocytes selectively and therefore may not totally reduce immunocompetence to bacterial infection. Many types of allotransplantation are now routinely successful because because of this drug (heart, heart-lung, liver, pancreas).

Dr. Achauer has shown that skin allografts on burned rats are not rejected when CyA is administered concurrently. While the immune response to alloantigens is suppressed, the response to bacterial infection is at least partially retained. The studies are investigating the interaction and combination of massive thermal injury and its immunosuppressive nature, the use of large amounts of allogeneic material required to cover a large burn, and the minimal effective CyA dose to successfully inhibit rejection while allowing enough immunocompetence to counter bacterial infection.

"Regeneration of Epidermis by Grafting Cultured Cells"
R01 GM33158-01A1 (Green, H.) Harvard University

Dr. Green and his collaborators have demonstrated it is possible to take a small sample of full thickness epidermis from a badly burned patient and grow the cells in culture so as to obtain enough cultured epithelium to cover large areas, and when this cultured epithelium is grafted to humans under clinical conditions, a high frequency of take is obtained. The current research efforts of Dr. Green and his collaborators include: a study of the multiplication and terminal differentiation of keratinocytes in culture in order to learn how to control those factors that govern graft behavior; a comparison of cultures derived from different donor sites and their effectiveness in grafting and the quality of the skin regenerated; and a study of methods of preserving cultures, both in frozen and metabolically active states.

Biophysics and Physiological Sciences Program

Program Organization and Professional Staff

Marvin Cassman, Ph.D.	Director
Frederick P. Ferguson, Ph.D.	Program Coordinator

Biophysics Section

James Cassatt, Ph.D.	Chief
John C. Norvell, Ph.D.	Program Administrator
Americo Rivera, Ph.D.	Program Administrator
Bert Shapiro, Ph.D.	Program Administrator
Helen Sunshine, Ph.D.	Program Administrator

Physiological Sciences Section

Lee Van Lenten, M.D.	Chief
Elizabeth O'Hern, Ph.D.	Program Administrator
William M. Taylor, Ph.D.	Program Administrator

Support Staff

Evelyn Laten	Secretary to the Director
Susan Bozak	Grants Clerk
Sharon Frazier	Grants Clerk
Bonnie French	Grants Clerk
Mary Beth Gallagher	Grants Clerk
Mollie Hilty	Grants Clerk
Wilma Kline	Grants Clerk
Judith Wilson	Grants Clerk

MINORITY ACCESS TO RESEARCH CAREERS PROGRAM

OBJECTIVES

The Minority Access to Research Careers (MARC) Program was formally established in 1975, and is a unique research training support program administered by the National Institute of General Medical Sciences in collaboration with other institutes of the National Institutes of Health. The primary goals of the MARC Program are to increase the number and capabilities of minority scientists engaged in biomedical research and to strengthen science curricula and research opportunities at minority institutions in order to prepare students for careers in biomedical research. To achieve these goals, the MARC Program uses both institutional research training grants and individual fellowships. These are: the MARC Faculty Fellowship, the MARC Visiting Scientist Award, the MARC Honors Undergraduate Research Training Grant, and the MARC Predoctoral Fellowship.

PROGRAM COMPONENTS

The MARC Faculty Fellowship, the first of the four MARC support mechanisms developed, provides opportunities for advanced research training for selected faculty members of 4-year colleges, universities, and health professional schools in which student enrollments are drawn substantially from ethnic minority groups (such as American Indians, Blacks, Hispanics, and Pacific Islanders). These institutions may nominate faculty members for MARC fellowships for a period of up to 3 years of advanced study and research training, either as candidates for the Ph.D. degree or as investigators obtaining postdoctoral research training in the biomedical sciences. Faculty Fellows are expected to return to the nominating institutions at the completion of their training.

MARC Faculty Fellows are selected on a competitive basis. Evaluation is based upon the applicant's qualifications and potential for research and training, as evidenced by academic records, reference reports, and publications, as well as by the research training proposal, the proposed training situation (i.e., the sponsor's and the institution's training facilities and staff), and other relevant information.

The second support mechanism, the MARC Visiting Scientist Award, provides financial support for outstanding scientist-teachers to serve as visiting scientists at 4-year colleges, universities, and health professional schools where student enrollments are drawn substantially from ethnic groups. The primary intent of this award is to help strengthen research and teaching programs in the biomedical sciences for the benefit of students and faculty at these institutions by allowing them to draw upon the special talents of expert scientists from other institutions. Reciprocal benefits should accrue to the Visiting Scientist through the added experience gained by his or her involvement in innovative science education and research development programs.

The proposal must include arrangements for the Visiting Scientist to reside in the campus community and to participate fully in programs of teaching, development of research, and/or counseling, as outlined by the institution. Evidence of negotiations and of some tentative agreement between the applicant institution and the scientist-teacher nominated to serve as the MARC Visiting Scientist

should be provided as a part of the application. The individuals nominated should be recognized scholars and leaders in the biomedical sciences. Proposals may request support for periods ranging from one academic quarter to a maximum of 1 year. Stipends are determined on an individual basis, according to the nominee's current salary or other possible source of stipend support for the proposed period in residence.

The MARC Honors Undergraduate Research Training Grant, the third mechanism of support, was initiated at the suggestion of Congress and Institute consultants and staff as a means of emphasizing the value and importance of providing biomedical research training at the undergraduate level in minority institutions. The objectives of the program are: to increase the number of well-prepared students who can compete successfully for entry into graduate programs leading to the Ph.D. degree in the biomedical sciences, to help develop a strong undergraduate science curriculum, and to stimulate an interest in undergraduate research as preparation for graduate study in the biomedical sciences.

Training support is offered to carefully selected undergraduate honors students at 4-year colleges, universities, and health professional schools in which student enrollments are drawn substantially from ethnic minority groups. Each institutional grant is awarded for a maximum period of 5 years. These honors programs for third- and fourth-year students should be designed to improve significantly the research training capabilities of the minority institutions. Applications for support should provide information regarding proposed mechanisms to augment and improve the science curricula, strengthen the faculty, and improve laboratory facilities. Funds are available for research equipment and supplies essential to the program, stipends, tuition, fees, and limited travel costs for the trainees. Arrangements for special training at universities and laboratories other than the grantee institution are judged essential and should be described in the application.

The fourth component of the program is the MARC Predoctoral Fellowship, which provides support for research training leading to the Ph.D. degree in the biomedical sciences for selected students who are graduates of the MARC Honors Undergraduate Research Training Program. It is expected that such training will be conducted in graduate-degree programs of the highest quality. Support is not available to individuals enrolled in medical or other professional schools, unless they are enrolled in a combined degree (M.D.-Ph.D.) program. Awards are conditional upon acceptance into a specified doctoral (Ph.D.) degree program in biomedical research.

MARC Predoctoral Fellows are selected on a highly competitive basis. A maximum of 5 years of support may be recommended, based on the merit of the application and evidence of satisfactory progress in the doctoral program in which a successful applicant is enrolled. The award provides an annual stipend to the student and funds to help defray the expenses of tuition, fees, and supplies.

ORGANIZATION AND STAFFING

The MARC Program is administered by its Director, Mr. Elward Bynum, and one professional staff member, Mrs. Dolores L. Lowery, as program administrator.

In addition, the senior consultant to the Institute, Dr. Geraldine P. Woods, provides ongoing assessment and evaluation of the program.

HIGHLIGHTS

Since the inception of the MARC Faculty Fellowship Program in 1972, approximately 444 applications for MARC fellowship support have been received, and 193 individuals have received support under the program. These individuals include 115 predoctoral and 78 postdoctoral award recipients. Of the 115 predoctoral fellows, 97 have now received the Ph.D. degree; of the 78 postdoctoral fellows, 69 have completed their training. Overall, 89 percent of those who have completed the program have returned either to the original home institution or to another minority institution.

The research training sites have included 70 universities, research laboratories, and federal institutions in 32 states and 5 foreign countries. The home institutions are broadly representative, including 68 universities in 23 states and the Commonwealth of Puerto Rico.

There has been little change in the level of support for the Visiting Scientist Award Program. Only 9 awards have been provided, based on 16 applications reviewed. The requirement for considerable advance planning by the host institution and the prospective Visiting Scientist may have tempered a fuller usage of this award. However, concerted efforts to advertise and encourage applications for the Visiting Scientist Award are being made by program staff. The MARC Review Committee and the National Advisory General Medical Sciences Council strongly support continuation of this award mechanism.

To date, 258 applications for the Honors Undergraduate Research Training Grant Program have been received, of which 157 have been approved; 53 minority institutions have received awards, 5 of which were second 5-year awards. In 1984, there were 207 graduates. In 1985, the program had 192 graduates.

In January 1981, the National Institute of General Medical Sciences invited applications for individual National Research Service Award MARC Predoctoral Fellowships to initiate this component of the MARC Program. Sixteen applications were received, of which 14 were approved and 12 funded. In 1982, the Institute received 29 applications, of which 19 were approved and 11 funded. In 1983, 42 applications were received, 25 approved and 11 funded. In 1984, 33 were received, 25 approved, and 17 funded. In 1985, 18 were received, 13 were approved, and 8 were funded.

EVALUATION

Assessment of the improvements and accomplishments of the various components of the MARC Program is a continuing effort of program staff. The scientific community, the MARC Review Committee, the National Advisory General Medical Sciences Council, and special Institute consultants cooperate in this effort through continued interaction with program and other NIGMS staff.

Significant accomplishments of the program include the following:

The "honors" concept of the MARC Program has encouraged institutions that are recipients of awards to establish a different and more rigorous program for students in the basic sciences. Many have incorporated an honors thesis requirement for graduating MARC students.

The provision of funds for seminars and lectures, release time for faculty members, up-to-date equipment and supplies, as well as the addition of full-time and part-time faculty in disciplines not formerly represented have all contributed to academic enrichment. In addition, specialized courses have been set up to prepare students to take standardized tests--an area where minority students have traditionally been weak.

The scientific activities of grantee departments have been strengthened, and as a result, it has become more common for MARC faculty members and trainees to be invited to present papers at local, regional, and national scientific meetings. We are beginning to see a larger number of students serving as coauthors of scientific papers.

The MARC Program, with its emphasis on research, has produced a new awareness of the possibilities available in biomedical research among students historically motivated towards careers in medicine, dentistry, and education. A higher percentage of the trainees are entering graduate schools.

The fact that MARC students are performing well in graduate schools has provided an increased number of "role models" for minority students, and has resulted in a significant increase in the number of science majors at minority institutions having MARC awards.

We are now seeing a number of the graduates of the Honors Undergraduate Research Training Grant Program who were supported in the first years of that award preparing to receive the doctorate in 1985. From information available, they have done extremely well and are viable candidates for postdoctoral support at prestigious institutions.

SUPPLEMENTAL ACTIVITIES

The Fourth MARC Scholars Conference and Program Directors Meeting was held on October 1 through 3, 1984 at the National 4-H Center in Chevy Chase, Maryland. One hundred and thirty-four MARC-supported college seniors attended the conference. Many had prepared oral and poster presentations on their research and spent much of the 3 days meeting with graduate school faculty and scientist representatives of the National Institutes of Health to discuss opportunities in graduate education. They received information on the graduate school application process, the availability of financial aid, and expectations regarding the Graduate Record Examination. As part of the conference, participants attended scientific seminars and toured selected NIH research laboratories. They also heard welcoming remarks by Mr. Elward Bynum and a keynote address by Dr. Bertram Fraser-Reid, Professor, Department of Chemistry, Duke University. Dr. Michael Potter, Chief, Laboratory of Genetics in the Division of Cancer Biology and Diagnosis, National Cancer Institute, addressed the scholars at lunch on the first day; Dr. Robert Barker, Provost, Cornell University, spoke at a banquet on the second evening.

Also held on the second day was a scientific seminar, with presentations on the state of the art in various research areas given by Dr. Paul St. John, Laboratory of Neurophysiology, National Institute of Neurological and Communicative Disorders and Stroke; and Dr. John Dement, Health and Safety Manager, National Institute of Environmental Health Sciences.

On the final day of the conference, a concurrent MARC Program Directors Meeting was held to provide a forum for the improvement of program planning and coordination of the MARC Honors Undergraduate Research Training Grants. Discussions centered on a number of issues such as requirements for renewal applications, requirements for predoctoral applications, training for faculty, summer research, training programs, the impact of MARC on the institutional setting, the use of the visiting scientist program, pre-MARC student activity, and grants management issues affecting each of the 50 MARC Program Directors in attendance.

The annual meeting of the American Society for Microbiology held on March 5-9, 1984 in Las Vegas, highlighted on its agenda a round table discussion on MARC Program activities. The directors of the program along with several MARC Program faculty from the MARC institutions discussed various aspects of the scope and impact of the programs. Questions generated in this forum centered on the summer research components, with many attendees expressing an interest in having MARC honor students in their laboratories.

Plans are underway to present the program at the annual meetings of the Federation of American Societies for Experimental Biology and the American Society of Biological Chemists to encourage faculty at prestigious institutions to select MARC scholars as trainees.

Minority Access to Research Careers Program

Program Organization and Professional Staff

Elward Bynum	Director
Dolores Lowery	Program Administrator

Support Staff

Mary Staley	Grants Clerk
Martha A. Brown	Clerk-Typist